


```

|||||
1 MetPheArgLeuGlnPheArgLeuPheProPheLeuArgThrAlaMetI 17
51 CATCCGTTGACCGCCCTGCTCAAAAGCTTCCTCCCTGCTGCTTCCT 100
17 stlleuLeuThrAlaLeuLeuLysSerLeuLeuProLeuSerC 34
101 GTCGACACGCTGGGAAACCGGCTGACATCTGGCTTTACCTTTA 150
34 ysleuHisThrleuGlyAsnArgleuGlyHisleuAlaPheThrleu 50
151 AAGGAAGACCGCGCGCATCTGCGCAATATGCGGACGCGGTTGAA 200
51 LysGluAspArgAlaArgIleValAlaAsnMetArgGlnAlaGlyMetAs 67
201 CCCGACACGACGAGCTGAAAGCGCTTTTGGGAAAGCGGAAATGGG 250
67 nProAspProLysThrValLysAlaValaPheAlaGluThrAlaLysGlyG 84
251 GTTTGAACTTGCCCGCGTTTTCAAAAGAACCGAAGACATCGAACA 300
84 LysleuGluLeuAlaProAlaPhePheArgLysProGluAspIleGluThr 100
301 ATCTTCAAGCGGTACACGCTGGGAAACGCTGACAGACGCTTGACAA 350
101 MetPheLysAlaValHisGlyThrGluHisValGlnGlnAlaLeuAspLys 117
351 GGGGCAAGGGCTGCTGTTTCATCAGCGCGCATGCGGACGCTGATTTG 400
117 shIsGluGlyLeuLeuPheIleThrProHisIleGlySerLysAspLeuG 134
401 GGGGACGCTACATCAGACGACGCTTCGCTTCACCTGACCGCATGAC 450
134 LysGlyArgThrIleSerGlnGlnLeuProPheProLeuThrAlaMetLys 150
451 AAGCGCGCAAAATCAAGCGCATAGCAAAATCATGCAAGCGGCGGACGT 500
151 LysProProLysIleLysAlaIleLysPylMetGlnAlaGlyArgVal 167
501 GCGGCGCAAGGCAAAACCGCGCGCGCATACAGCGGTCAAAAGAA 550
167 LArgGlyLysGlyLysThrAlaProThrSerIleGlnGlyValLysGlnI 184
551 TCATCAAGCGCTGGCGCGCGGCGGACGACCATCATCTCCGCCGACAC 600
184 IeIleLysAlaLeuArgSerLeuGlnAlaThrIleValLeuProAspHis 200
601 GTCCCTTCCGACGAGAGCGCG...GCGGCTGGCGGATTTTTCGG 647
201 ValProSerProGlnGluGlyLysIleLysValItrPValAspPhePheG 217
648 CAACCTCGATACACCATGACACTGGCGCAAAATGGGACACATCAAG 697
217 LysProAlaLysThrMetThrLeuAlaIleLysLeuAlaHisValLysG 234
698 GCGTAAACCGCTGTTTTCGCTGCGGAAACCGCTCCCGCGACGACAGG 747
234 LysAlaLysThrLeuPhePheCysGlyLysGluArgLeuProGlyLysIle 250
748 TTCGATGTCACATCCGCGCGCGTCAAGGGAATTTGAAGCGCAAGG 797
251 PheAspLeuHisIleArgProValGlnGlyLysLeuAsnGlnLysAsp 267
798 CCAGATGCGCGCGGTTCACACCGCAATACGAATATTGGATACCGCT 847
267 ahIsAspAlaIleValaPheAsnArgAsnAlaGlyLysTrpIleArgLysP 284
848 TTCGACGACGATATCTGTTATGTACAAACGCTATAAAGCGCG 891
284 heProThrGlnThrLeuPheMetLysAsnArgLysMetPro 298
seq_name: sp_bacteria:Q9EYMO

```

```

seq documentation_block:
ID Q9EYMO PRELIMINARY; PRT; 337 AA.
AC Q9EYMO;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE HTRB PROTEIN.
OS Acinetobacter sp. BD413.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=104611;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BD413;
RX MEDLINE=21318707; PubMed=11425734;
RA Friedrich A., Hartsch T., Averhoff B.;
RT "Natural transformation in Mesophilic and Thermophilic Bacteria:
RT Identification and Characterization of Novel, Closely Related
RT Competence Genes in Acinetobacter sp. Strain BD413 and Thermus
RT thermophilus HB27."
RL Appl. Environ. Microbiol. 67:3140-3148(2001).
DR EMBL: AF320001; AAG34711.1; -.
DR InterPro: IPR000515; BPD.transp.
DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER: UNKNOWN.1.
SQ SEQUENCE 337 AA; 38795 MW; 6E998AFEDICEB140_CRC64;

```

```

alignment_scores:
Quality: 288.00 Length: 278
Ratio: 1.665 Gaps: 7
Percent Similarity: 62.230 Percent Identity: 29.496

```

alignment_block:

US-09-303-518D-571 x Q9EYMO ..

Align seg 1/1 to: Q9EYMO from: 1 to: 337

```

55 CTGTGACCGCCGCTGCTCAAAATGCTTCGCTGCTGCTTCCTGCT 104
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
31 MetLeuAlaGlyLeuValaAsnThrLeuLysLeuThrLysThrSer..... 45
105 GCACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTACCTTTAAAG 154
:::||||:|||||:|||||:|||||:|||||:|||||:
46 ArgThrIleGlnLeuAsnLeuLysIleThrLeuProTyrLeuAsnAsp 62
155 AAGACGCGCGCGCATGCTGCCAATATGGCGGACGCGGTTGAACCC 204
62 GlnAlaArgGluArgIleIleGlnAsn..... 70
205 GACACGACGACGCTCAAAACCGCTTTTGGCGAAACGCAAAATGCGGTT 254
|||||:|||||:|||||:|||||:|||||:|||||:
71 .....AlaIleArgAsnGlnLeuThrSerLys...Me 80
255 GGAACCTTGCCCGCGCTTTTCAAAAACCGGAGACATGGAACAAATGT 304
:::||||:|||||:|||||:|||||:|||||:|||||:
80 GlnPhePheSerIleThrPheLysSerAsnGlnLysAsnIleSerArgI 97
305 TCAAGGCGGTACACGCGTGGGACAGTGAGGAGCGCTTTGGCAAGGCG 354
:::|||||:|||||:|||||:|||||:|||||:|||||:
97 LeuHisValaGluGlyAlaAspPhePheHisGlnAlaIleAlaGluAsn 404
:::||||:|||||:|||||:|||||:|||||:|||||:
114 LysGlyIleValLeuIleValaProHisPheGlyThrThrGluIleMetAs 130
405 ACCCTACATGACGACAGCTTCGCTTCACACCTGACGACGCGCATATGAC 454
:::||||:|||||:|||||:|||||:|||||:|||||:
130 AlaItrPheCysAlaGlnThr.....AspMetThrIleLeuLysLysP 145
455 GCGCGAAATCAAGCGCATAGCAAAATCATGCAAGCGGCGGACGCGGCG 504
||| ||| |||:|||||:|||||:|||||:|||||:
145 roValLysAspLysAspAlaAspArgPheValaArgGlnAlaArgSerArg 161

```

```

505 GGCAAGGCAAAACCGCCGACGCGCTACAGGGGTCAAAACATCAT 554
    :::::::::: :::::::::: ::::::::::
162 GluGlnAlaValLeuValProThrAspGluValArgGlnIlePh 178
555 CAAGGCGCCGCGCGGCGAGGCAACCATCATCTGCCGCGACGCGC 604
    :::::::::: :::::::::: ::::::::::
178 eLysAlaLeuLysGlnGlyThrThrValIleLeuProAspHleTrp 195
605 CTTCTCCGAGGAGGAGCGCGCGCTGTGGCGGATTTTTCGCAACCT 654
    :::::::::: :::::::::: ::::::::::
195 roAsnVal.....GlyGlyGluMetValAspTyrPheGlyLeuPro 208
655 GCATACACCATGACACTGGCGGCAAAATTGGCACACGTCACAGGCGTAA 704
    :::::::::: :::::::::: ::::::::::
209 LeuAlaSerSerAsnLeuSerAlaLysLeuIleGlnTyrThrLysAla 225
705 AACCGCTGTTTCTGTGCGAGCGCTGCCGCGACAGACGCTCGCTGT 754
    :::::::::: :::::::::: ::::::::::
225 sAlaLeuPheLeuTyrAlaIleArg...AsnGluAsnHisGlyPheAsp 241
755 TGCACATCCGCGCCGCTCCAGGCGGAATTGAACGCAACAAAGCCGCGAT 804
    :::::::::: :::::::::: ::::::::::
241 ePheIleLeuProMetAspProAlaIleTyrGlnGlyThrAspHisasp 257
805 GCGCGCC...GTGTTACACCGCATACCGAATTTGGATACGCCGTTTTC 851
    :::::::::: :::::::::: ::::::::::
258 GlyThrLeuValIleHisHisAlaIleGluAspLeuIleArgLysTyr 274
852 GACGCGATCTGTTTATGTACACCGCATATAA 885
274 oAspHisTyrHisTyrPseTyrTyrLysArgPheLys 285

seq_name: sp_bacteriap:Q917B5
seq_documentation_block:
ID Q917B5 PRELIMINARY; PRT; 295 AA.
AC Q917B5;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE PROBABLE 2-OH-LANUROYLTRANSFERASE.
GN PA0011.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.B., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Birkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laribi K., Jim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL: AE004441; AAC03401.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 295 AA; 33001 MW; D6D1710EC6AD42FF CMC64;

```

alignment_scores:

Quality:	266.50	Length:	285
Ratio:	1.514	Gaps:	7
Percent Similarity:	61.754	Percent Identity:	26.667

alignment_block:

US-09-303-518d-571 x Q917B5 ..

Align seq 1/1 to: Q917B5 from: 1 to: 295

```

55 CTTGTACCGCCCTGCTCAAAATGCTCTCCCTGCTGCTTTCCTGCT 104
    :::::::::: :::::::::: ::::::::::
8 LeuValValGlyAlaLeuValLeuPheAlaLeuLeuProTrrArgAlaVa 24
105 GCACACGCTGGAAACCGGCTCGACATCTGCGCTTACCTTTTAAAG 154
    :::::::::: :::::::::: ::::::::::
24 IginGlyValGlyAlaGlyIleLeuTyrPheLeuMetTrrPlySleu.....P 39
155 AAGACCGCGCGCGCATGTCGCCAATATGCGGAGCGGCGGTTTGAACCC 204
    :::::::::: :::::::::: ::::::::::
39 roAsnArgSerArgGluValValArgIleAsnLeuSerLysScyPhePro 55
205 .....GACACGCAAGCGTCAAAAGCGCTTTTTCGGGAAAGCGCAA 245
56 GluLeuSerGluThrGluLeuGluLysLeuValGlyGlnSerLeuMetAs 72
246 ATGCGGT.....TTGAACTGCCCCCGGCTTTTCAAAAACCGG 286
    :::::::::: :::::::::: ::::::::::
72 pIleGlyArgThrLeuThrGluSerAlaCysAlaTrrPleTrrProProG 89
287 AAGACATCGAACAATGTTCAAGCGGTACACGCTGGGACACGTGCAG 336
    :::::::::: :::::::::: ::::::::::
89 LulysSerLeuArgTyrIleArgGluValGluGlyMetGluValLeuGlu 105
337 CAGCGTTTGACAAAGGCGGAGGCTGCTGTTCATACAGCGCGCACATCG 386
    :::::::::: :::::::::: ::::::::::
106 GluAlaLeuAlaSerGlyAspGlyLeuValGlyIleThrSerHisLeuG 122
387 CAGCTACGAT...TTGGCGGAGCTGTACATCAGCCAGCAAGCTTCCTCC 433
    :::::::::: :::::::::: ::::::::::
122 yAsnTrrpGluValLeuAsnHisPheTyrCysSerTyrAlaLysProIle. 138
434 ACCTGACCGCATGTACAGCCGCGGAAATCAAGAGCATAGCAAAATC 483
    :::::::::: :::::::::: ::::::::::
139 .....IlePheTyrArgProProLysLeuLysAlaValAspGluLeu 152
484 ATGACAGCGGCGAGGCTGCGCGGCAAAACCGCCACCGCGCAT 533
    :::::::::: :::::::::: ::::::::::
153 LeuLysLysGlnArgValGlnLeuGlyAsnArgValAlaProSerThrP 169
534 ACAAGGCGTCAAAACAAATCATCAAGGCTGCGCGCGGCGGAGCAACCA 583
    :::::::::: :::::::::: ::::::::::
169 oGluGlyIleLeuSerValIleLysGluValLysGlyGlyCysValG 186
584 TCATCTGCGCGACACGCTCTCTCCGACGAAGCGCGCGCTGTG 633
    :::::::::: :::::::::: ::::::::::
186 LylIleProAlaAsp.....ProGluProAlaArgTrrAlaGlyLeuPhe 200
634 GCGGATTTTTCGCAAACTGCTACACCATGACACTGCGCGCAAAAT 683
    :::::::::: :::::::::: ::::::::::
201 ValProTyrLeuGlyTrrThrAlaLeuIleSerLysPheValProGln 217
684 GGCACACGTCAAAAGCGTGAACAACTGTTTCTGCTCGGAAGCGCTGC 733
    :::::::::: :::::::::: ::::::::::
217 uLeuSerArgGlyLysAlaArgGlyValPhePheHisAlaValArgLeuP 234
734 CCGACGCAAGGCGTGTGTCATCGCGCGCGCGCGCAAGGGAATG 783
    :::::::::: :::::::::: ::::::::::
234 roAspGlySerGlyTyrLysValIleLeuGluAlaAlaProAlaAspMet 250
784 AAGCGCAACAAAGCCGAC...GATGCGCGGTGTTTCAACCGCAATACGA 830
    :::::::::: :::::::::: ::::::::::
251 TyrAspLysAspLeuGluValSerValAlaAlaMetSerArgLysLeuAl 267
831 ATATTGATACGCGGTTTTCGACGACAGTATCTGTTATGTACACCGCT 880
    :::::::::: :::::::::: ::::::::::
267 aLysTyrValArgAlaTyrProSerGlnTyrMetTrrPseMetLysArgP 284
881 ATAAA 885
284 helys 285

```


RA MEDLINE=98340543; PubMed=9675892;
 RA Dekkers L.C., van der Bij A.J., Mulders H.M., Phoele C.C.,
 RA Wentwood A.R., Glandorf D.C.M., Wijffelman C.A., Lugtenberg B.J.J.,
 RT Role of the O-antigen of lipopolysaccharide, and possible roles of
 RT growth rate and NADH:ubiquinone oxidoreductase (nuo) in competitive
 RT tomato root-tip colonization by *Pseudomonas fluorescens* WCS365.⁺
 RL Mol. Plant Microbe Interact. 11:763-771(1998).
 DR EMBL: Y14566; CAY/4897.1; -
 FT NON_TER 253 253
 SO SEQUENCE 253 AA; 27634 MW; F12E003B10C8223D CRC64;

alignment_scores:

Quality: 226.00 Length: 248
 Ratio: 1.477 Gaps: 8
 Percent Similarity: 61.694 Percent Identity: 29.435

alignment_block:

US-09-303-518D-571 x 033437 ..

Align seg 1/1 to: 033437 from: 1 to: 253

```

55 CTGTTGACCGCCCTGTCAAATGCTCTGCTGCTGCTTCTCTCTCT 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 LeuLeuVal1G1Val1LeuArgLeuPheAlaLeuLeuProTyrAlaVala 24
105 GCACACGCTGGGAACCGCGCTCGACATCTGCGCTTACCTTTAAAG 154
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 IG1AlaVal1G1Ser1Ala1IleG1YTrpHemTrpLysLeu.....P 39
155 AAGACGCGCGCGCATGCTGCCCAATGATGCGGACGCGGTTTGAACCC 204
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
39 roAsnArgSer1Arg1spValValArg1ValAsn1LeuAla1LysCysPhe 55
205 GAC.....ACGACGACGCTCAAGCGCTTTTGCGGA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 GluMetAspProAlaGluArg1GluArg1LeuVal1G1GlnSer1LeuLys 72
237 AAGCGCAAAATGCGCTTGGACCTGCGCCCGCGCTTTTCAAAAAACGG 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 pIleG1LysSer1LeuThrG1Ser1AlaCys1AlaTrp1IleTrpPro1Ag 89
287 AAGACATCGCAAAATGTTCAAGCGGTCACACGCTGGGACACGTCGAG 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
89 lArgSer1Lea1spLeuVal1Arg1GluVal1G1Lys1LeuG1Val1LeuLys 105
337 CAGGCTTGGACAGGCGGACGCGCTGTCATCAGCCGCGCACATCGG 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 GluAlaLeuAlaSer1Lys1GluVal1Val1G1Lys1Leu1Ser1His1LeuG1 122
387 CAGCTACGAT...TTGGCGGACGCTACATCAGCCGACGCTTCCGTTCC 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
122 yAsnTrpG1Lys1Val1LeuAsn1His1Phe1Lys1Ser1G1Lys1Pro1Le 138
434 ACCTGACCGCCATGTCACAGCGCGCAAAATCAAGCGCATAGCAAAATC 483
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 .....lIlePhe1YrArg1Pro1Lys1Leu1Val1Ala1Val1sp1Leu 152
484 ATGACGCGCGGCGGCTGCGCGGCAAAAGCAAAAGCCGCGCACGCGAT 533
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 LeuArg1Lys1GlnArg1Val1GlnLeu1Gln1Lys1Val1Ala1Ala1Ser1Thr1Lys 169
534 ACAAGGGGTCAAAATCATCATCAAGGCGCTGCGCGGCGGACGAGCAACA 583
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 sgl1Glu1Lys1LeuSer1Val1IleLys1GluVal1Arg1Lys1G1GlnVal1G 186
584 TCATCTGCGCGGACGACGTCCTTCTCCGACGAGAAAGCGCGCGGTGAG 633
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 lYle1Pro1Ala1sp.....ProGlu1Pro1Ala1Gln1Ser1Ala1G1Lys1IlePhe 200
634 GCGGATTTTTCGCAAACTGATACCATCATGACATGACGCGGCAAAA... 681
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 Val1Pro1Phe1Ala1ThrG1Ala1Leu1ThrSer1Lys1Phe1Val1Pro1Asn1He 217

```

682 .TTGGCAGACGTCAAAGCGCGTGAACCCGTTTCTGCTGCGAACC 730
 217 CleuAla1G1Lys1Lys1Ala1Val...GlyVal1Phe1Leu1His1Ala1Leu1Arg1 233
 731 TGCCCGACGAGACAGGCTTCT.....GTGTCACATCCGCGCC 768
 233 euPro1Asp1GlySer1Gly1Lys1Val1Ile1Leu1Glu1Ala1Pro 247

seq_name: sp_bacteriap:Q9HYZ8

seq_documentation_block:

ID Q9HYZ8 PRELIMINARY; PRT; 312 AA.
 AC Q9HYZ8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE PROBABLE LAUROYL ACYLTRANSFERASE.
 GN PA3242.
 OS *Pseudomonas aeruginosa*.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garner R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Lim R.M.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbén I.T.,
 RA Smith K.A., Spencer D.H., Wong G.K.S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL: AE004747; AAC0630.1; -
 KW transferase; Acyltransferase; Complete proteome.
 SQ SEQUENCE 312 AA; 36063 MW; 0EDF771D75F3C186 CRC64;

alignment_scores:

Quality: 222.50 Length: 311
 Ratio: 1.236 Gaps: 11
 Percent Similarity: 57.878 Percent Identity: 26.045

alignment_block:

US-09-303-518D-571 x Q9HYZ8 ..

Align seg 1/1 to: Q9HYZ8 from: 1 to: 312

```

33 TTTGCGAAGCGCCATGACATCCGTGACCGCCCT..... 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 Phe1Ser1Ala1Phe1Leu1His1Pro1Arg1YrTrp1Pro1Leu1Trp1Heg1Lys 22
69 .....GTCAAATGCTCTGCTGCTGCTGCTGCTTCC 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22 uG1Leu1Leu1Trp1Leu1Val1Val1Gln1Leu1Pro1YrTrp1Val1Leu1Met1 39
100 TGCTGACACAGCTGGGAACCGGCTGCGACATGCGCGTTTACCTTT 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
39 eu.....Gly1Arg1Glu1Leu1G1Ala1Leu1Met1Lys 48
150 AAAGAGACCGCGCGCGCATGCTGCGCAATATGCGGACGCGGCTTGA 199
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
49 .....Arg1Leu1Val1G1Ser1Arg1Arg1Glu1Leu1Ala1Arg1Asn1LeuG1 63
200 A.....CCCGACAGCAGACGCTCAAA 222
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 uLeu1Cys1Phe1Pro1Lys1Ser1Pro1Ala1Glu1Arg1Glu. Arg1Leu1Lys 79
223 GCGGTTTTCGGAAGCGCAAAATGCGGTTTGAACCTTGCCTCCCGCGGT 272
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

80 GluAspRheA1aSerSerg1y1ealaphRheglumetalMetSert 96
273 TTTCAAAAAAGCGAGACATCGAATGTTCAAGGCGTACAGCGCT 322
    |||
96 pTTP...TTProlYalAArgLeuAlaArgLeuAlaHisIleGlyL 112
323 GGGACACATGCGACGAGGCTTTGGACAGGCGGAAAGGCTGTTGATC 372
    |||
112 euDHisLeuArgLysIleAlaGlnIleGlnIleGlyL 128
373 AGCGCCCATGCGGACGATTCGATTTGGCGGACCGTACATAGCCAGCA 422
    |||
129 AlaLeuHisRheThrLeuGluIleGlyAlaAlaLeuLeuGlyL 144
423 GCTTCGCTTCAC...CTGACCGCATATGACAGCGCGCAAAATCAAG 469
    |||
145 .....ValHisThrIleAspGlyMetGlyArgGluHisAspAsp 159
470 CGATAGACAAATCATGACAGGCGGAGGCTCGCGCAAGGCAAAAGC 519
    |||
159 alRheAspTyrValGlnArgArgGlyArgGluArgHisIleAspAla 175
520 GCGCCACCGCATACAGGCTCAACAAATCATACAGGCGCTGCGCGC 569
    |||
176 ThrAlaIleGluArgGluAspValArgSerMetLeuLysValLeuArg 192
570 GGGCGAGGCAACCATCATCCCTGCGGACGCTCTCCGACAGAG 619
    |||
192 yglArgAlaIleTrrPyrAlaProAsp.....GlnAspT 204
620 GGGCG.....GGCGTGGCGGATTTTTCGCAAACTGCATAC 660
    |||
204 yrcIyAlaLysGlnSerLeuRheValProLeuRheGlyIleProAla 220
661 ACATGACACTGCGGCAAAATTTGGACACGTCGAAGGCGGAAACCT 710
    |||
221 ThrValThrAlaThrThrLysRheAlaArgLeuGlyArgAlaValLe 237
711 GTTTTCTGCTGCGAAGCGCTGCGGACGAGGCTTCTGTTGACACA 760
    |||
237 uProRheThrGlnSerArgLeuAlaAspGlySerGlyTyrArgLeuPrt 254
761 TCCGCGCGGCTGCAAGGGAATGGAAGGC...AACAAAGCCGACATGC 807
    |||
254 LeuIleProGluGluAspRheProGlyGluSerGluAlaAspLys 270
808 GCGGTGTCAACCGCAATACGATATGATGAGCGGCTTTCGACAGCA 857
    |||
271 LeuArgIleAsnGlnTrrValGlnArgCysValArgGlnIleProGlu 287
858 GTATCTGTTATGTACAAACGCTATTAACG 888
    |||
287 nTyrLeuTrrAlaHisArgArgRheLysThr 297
seq_name: sp_bacteriar_09JYV2
seq_documentation_block:
AC 09JYV2 PRELIMINARY; PRT; 289 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GN HTRB/MSB8 FAMILY PROTEIN.
OS Neisseria meningitidis (serogroup B).
OX Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
RN NCBI_TaxID=491;
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RA MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,

```

```

RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Cleeke A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Colton M.D., Ullrich T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizarro M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RL Science 287:1809-1815(2000).
DR EMBL: AE002491; AAF4179.1; -.
DR TIGR: NMB1418; -.
KW Complete proteome.
SQ SEQUENCE 289 AA; 33843 MW; 3BF100F050576512 CRC64;

alignment_scores:
Quality: 204.00 Length: 289
Ratio: 1.207 Gaps: 12
Percent Similarity: 58.478 Percent Identity: 27.336

alignment_block:
us-09-303-518d-571 x 09JYV2 ..
Align seg 1/1 to: 09JYV2 from: 1 to: 289

67 CTGCTCAATGCTCTCCGCTGCTGCTTCCGTCGACACGCTGG 116
    |||
7 ValLeuTyrValLeuGlnPheLeuProHeaLeuLeuHisLysIleAl 23
117 AAGCGGCTCGACATCGGCTTTCATCTTTTAAAGAGACGCGCGC 166
    |||
23 asPLeuThrGlyLeuLeuAla...TyrLeuLeuValLysProArgArg 39
167 GCATC.....GTGCGCAATATGCGGACGCGGCTTGAACCC 204
    |||
39 rglIleGlyIleAsnLeuAlaLysCysPheSergIuTrrSergIuL 55
205 GACACGACGCGGTC...AAGCGCTTTTGGCGGAAAGCGCAAAATGCGG 251
    |||
56 LysArgLysThrValLeuLysGlnHisRheLysHisMetAlaLysLeu 72
252 TTGGAACTTGCSCCGCTTTTCAAAAAACCGAAGACATCGAACA 301
    |||
72 tLeuGlnTyrGlyLeuTyrTrrPyrAlaProAlaGlyArgLeuLysS 89
302 TGTTCAAAGCGGTACACGCTGGAAACGTCGACAGGCTTGAACA 351
    |||
89 euValArg...TyrArgAsnLysHisTyrLeuAspRasAlaLeuAla 104
352 GCGGAAGGCTGCTGATCATGACGCGGACATGCGGACGATGTTGG 401
    |||
105 GlyLysValIleLeuLeuTyrProHisRheThrAlaRheLeuAl 121
402 CGAGCGCTACATGACGACGAGCTTCCACCTGACGCGCATGTACA 451
    |||
121 aValTyrAlaLeuAsnGlnAspIlePro.....LeuIleSerMetLys 136
452 AGCGCGCGAATCAAGCGATGACAAATCATGACGCGGCGGAGCA 501
    |||
136 erHisGlnLysAsnLysIleLeuAspGluGlnIleLysGlyArgAsn 152
502 GCGGCGCAAA.....GCCAAACCGCGCGCCGACGCGGATACA 536
    |||
153 ArgTyrHisAsnValRheLeuIleGlyArgThr.....Gln 164
537 AGGGGTCAAAATCATGACAGGCTTCCGCGGCGGCGAGCAACATCA 586
    |||
164 uGlyLeuArgAlaLeuValLysGlnPheArgLysSerSerAlaProPhe 181
587 TC...CTGCGCGACCGACGCTCTTCCGCGAGAGGCGCGC..... 624
    |||
181 euTyrLeuProAsp.....GlnAspPheLysArgAsnAsp 192

```



```

625 GGGGTGGGGGATTTTTCGGCAACCTGCATACCATGACACTGGC 674
      :::::|||||
193 SerValrPheValaIAspPheHeGlyIleGlnThrAlaThrIleThrGlyLe 209
      :::::|||||
675 GGCAAAATTTGGCAGCTCAAGGCGGTGAAGCCCTGTTTCTCTGGG 724
      :::::|||||
209 uSerAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleProValIleProVal 226
      :::::|||||
725 AACGCTGGCCCGCAGCAAGGCTTTCGTTGCATCCGCGCCCTGCAA 774
      :::::|||||
226 AlaArgIuAlaAspAsnThr...ValThrLeuHisPheTyProAlaIle 241
      :::::|||||
775 GGGGATTTAGCGGCAACAAGCCAC...GATCGCGCGCTTTCAGCG 821
      :::::|||||
242 LysSerPheProGlyIuAspAlaIysAlaAspAlaIleArgIuMetAsnThr 258
      :::::|||||
822 CAATACCGAATATTGGATACGCGCTTTCGAGCGAGATCTGTTATGT 871
      :::::|||||
258 gPheIleGluAspArgValaArgIuHisProGluGlnIleTyPheIlePhe 275
      :::::|||||
872 ACAACCGCTATAAAGC 888
      :::::|||||
275 IAspArgPheLysThr 280
      :::::|||||

```

seq_name: sp_bacteriap:Q9JTV4

seq_documentation_block:

```

ID Q9JTV4 PRELIMINARY; PRT; 289 AA.
AC Q9JTV4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE ACETYLTRANSFERASE.
GN NMA1630.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parhill J., Achtman M., James K.D., Bentley S.D., Church C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagers K., Leather S., Moutle S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis 22491."
RL Nature 404:502-506(2000).
DR EMBL: AL162756; CAB84858.1; -.
KW Transferase; Complete proteome.
SQ
SEQUENCE 289 AA; 33867 MW; 4FAE453B5A632CID CRC64;

```

alignment_scores:

```

Quality: 202.00 Length: 288
Ratio: 1.195 Gaps: 10
Percent Similarity: 58.681 Percent Identity: 26.389

```

alignment_block:

US-09-303-518D-571 x Q9JTV4

Align seg 1/1 to: Q9JTV4 from: 1 to: 289

```

67 CTGCTCAATGCTCTCCCTGCTGCTTCTGTCGACACGCTGGG 116
      :::::|||||
7 ValLeuTytrValleuGlnPheLeuProPheAlaLeuLeuHisLysLeuAl 23
      :::::|||||
117 AAACCGGCTCGACATCTGGCGTTTACCTTTAAAGGAAGACCGCGCGC 166
      :::::|||||
23 aAspLeuThrGlyLeuLeuAlaTytrLeuLeuValLysProArgArgArgI 40
      :::::|||||

```

seq_name: sp_bacteriap:Q9PH42

seq_documentation_block:

```

ID Q9PH42 PRELIMINARY; PRT; 279 AA.
AC Q9PH42;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.

```

```

167 GCATCGTGGCAATATGCGGAGGAGGTTTGAACCCGACGACGAGAG 216
      :::::|||||
40 IeGlyIuIleAsnIleAlaIleLysCysPheProGluTytrAspGlyLys 56
      :::::|||||
217 GTCAAGCCGTTT.....TTTGGGAAACGGCAAAATGCGGTTT 254
      :::::|||||
57 ArgYsThrValIleuLysGlnHisPheLysHisMetAlaLysLeuMetLe 73
      :::::|||||
255 GGAACCTGGCCCGCGTTTTCGAAAAAACCGGAAGACATGCAAAACATG 304
      :::::|||||
73 uGluTytrGlyLeuTytrTytrAlaProAlaGlyArgLeuLysSerLeu 90
      :::::|||||
305 TCAAGCGGTACAGCGCTGGGGAACAGCTGCAGGAGCTTTGCAAAAGGC 354
      :::::|||||
90 AlaArg...TytrArgAsnLysHisTytrLeuAspAspAlaIleAlaIleGly 105
      :::::|||||
355 GAAGCGCTGCTGTTTCATACGCGCGCACATGCGGACGTACGATTTGGGCG 404
      :::::|||||
106 GluTytrValIleIleIleuTytrProHisPheThrAlaPheGluMetAla 122
      :::::|||||
405 ACGGTACATCAGCGAGCTTCGTTCCAGCTGCAGCGCATGACAGC 454
      :::::|||||
122 ITrAlaLeuAsnGlnAspValPro.....LeuIleSerMetTytrSer 137
      :::::|||||
455 GCGGAAATATCAAGCGATAGCAAAATCATGCGAGCGGCGAGGCTGCGC 504
      :::::|||||
137 IsgIuLysAsnLysIleLeuAspGluGlnIleLeuLysGlyArgAsnArg 153
      :::::|||||
505 GGCAAA.....GGCAAAACCGCGCCACCGGCAATACAGG 539
      :::::|||||
154 TytrHisAsnValPheLeuIleGlyArgThr.....GluI 165
      :::::|||||
540 GGTCAAAACAAATCATCAAGCGCTGCGGCGGCGGAGGCAACCATCATC 588
      :::::|||||
165 YLeuArgAlaLeuValLysGlnPheArgLysSerSerAlaIlePheLeu 182
      :::::|||||
589 ..CTGCGCCAGCCAGCTCCCTTCTCCGAGAGAGGCGC.....GCG 627
      :::::|||||
182 YrLeuProAsp.....GlnAspPheGlyArgAsnAspSer 193
      :::::|||||
628 GTTGCGGCGGATTTTTCGCAAACTGCATACACATGACACGCGGCG 677
      :::::|||||
194 ValPheValaIAspPheGlyIleArgThrAlaThrIleThrGlyLeu 210
      :::::|||||
678 AAATTTGGCACAGCTCAAGGCGGTGAAGCCCTGTTTCTGTCGCACAC 727
      :::::|||||
210 rArgIleAlaIleAlaIleAlaIleAlaIleValIleProAlaIleProVala 227
      :::::|||||
728 GCCTTCGCGGAGGACAGGCTTTCGTTGCATCCGCGCTCCAAAGG 777
      :::::|||||
227 rGluAlaIAspAsnThr...ValThrLeuHisPheTytrProAlaIle 242
      :::::|||||
778 GAATGAAACGGCAACAAGCCAC...GATCGCGCGCTTTCAGACCGG 824
      :::::|||||
243 SerPheProSerGluAspAlaIleAlaIAspAlaIleArgIuMetAsnArgPh 259
      :::::|||||
825 TACGAAATATTGGATACGCGCTTTCGAGCGAGATCTGTTATGTACA 874
      :::::|||||
259 eIleGluGluArgValaArgIuHisProGluGlnIleTytrPheIlePhe 276
      :::::|||||
875 ACCGCTATAAAGC 888
      :::::|||||
276 YAspArgPheLysThr 280
      :::::|||||

```


GN XF0104.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 NCBI_TaxID=2371;
 RX MEDLINE=20365717; PubMed=10910347;
 RC STRAIN=9A5C;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Abencio M.,
 Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
 Coutinho L.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 Fraga J.S., Franca S.C., Franco M.C., Froime M., Furian L.R.,
 Gartner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 Ho P.L., Honnisch J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Matsukuma A.Y.,
 Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 Menck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Vitorello C.B.,
 Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 Pelixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
 Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 de Rosa V.E.Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A.Jr.,
 da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 de Souza A.P., Tenenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
 Zago M.A., Zatz M., Melandris J., Setubal J.C.;
 RA "The genome sequence of the plant pathogen *Xylella fastidiosa*."
 RL Nature 406:151-159(2000).
 DR EMBL: AE003864; AAF82917.1; -
 KW Complete proteome.
 SO SEQUENCE 279 AA; 32072 MW; 450EC409876A7791 CRC64;

alignment_scores:
 Quality: 192.00 Length: 270
 Ratio: 1.255 Gaps: 10
 Percent Similarity: 56.667 Percent Identity: 27.407

alignment_block:

US-09-303-518d-571 x Q9PH42

Align seg 1/1 to: Q9PH42 from: 1 to: 279

112 CTTGGGAAACGGCTCGGACATTCGCGTTTACCTTTAAAGGAGACCG 161
 15 LeuGlyAlaGValLlleglYAlaMeLaLeuAlaLeuValGlyThraGAr 31
 162 CGCGGCAATCGTCGCAATATCGCGGAGCGGTGAACCCCGACCG 211
 31 gCysAlaAlaGluValAsnLeuArgLeuGlyPheProThrHisAspGluA 48
 212 AGACGCTAAAGCCGTTTTCGGGAAACGCGCAAAATGC.....GGT 252
 48 lATrPArGLySLySLySLySLySLySLySLySLySLySLySLySLySLy 64
 253 TTG...GAACCTGGCCCCCGTTTTCAAAAAACCAGGAAACATGCAAC 299
 65 LeuPheGluPheAlaArgAlaIleTrrP.....GlySerIleAspVa 78
 300 AATGTTCAACCGGTACAC.....GGCTGGGAACACGTCGACGAGGCTT 343
 78 lLleTrrGAlaGlyIleHisValIGluGlyLeuGluHisLeuHisIleuG 95
 344 TGGACAAGGCGGAGGGGTGTCTTCATCAACGCGGCACATCGGACGCTAC 393

95 lnaTrrGluGlyArgGlyValLeuLeuValSerLeuHisPheMetThrLeu 111
 394 GATTGGGCGGAGCGTATCATACACCCAGACACCTTCCTTCACGTCACCG 443
 112 GluLleGlySLyArGLyLeuGlyCysAspTyrValPro.....LeuAla 126
 444 CATGTACAAACCCCGCAAAATCAAGGAGTAGCAAAATCATTCACGACGG 493
 126 yMeTTrArgArGlnHisArgAsnProValLeuGluTrpAlaIleValArg 143
 494 GCAAGGTGGCGGCAAAAGGCAAAACCGCGCCACCGCATACAAAGGGTTC 543
 143 lYArGLyLeuArg.....TyrAlaThrGlnIlePheHisLeuGluGlu 157
 544 AAACCAATATATCAAGCCCTTCGCGGCGGCGGCGGAGACATCATCTGTC 593
 158 ArgAlaThrIleLeuHisLeuLysArgGlyPheLeuTrpYAlaLar 174
 594 CGACACATTCCTTCCTTCGCGGAGGAGGCGGCGGCGGCGGAGATTTC 643
 174 oAspGlnAspMetArgGlyLysGlu.....ThrValPheAlaProPhe 189
 644 TCGGCAACCTGATACACATGACATGCGCGGCAAAATTCGACACGTC 693
 189 heGlyMetProAlaAlaThrIleThrAlaThrHisGlnLeuAlaArgLeu 205
 694 AAAGCGGTGAACCCCTGTTTTCGTCGCGGAGCCCTGCCC..... 735
 206 ThrGly.....CysAlaValValProTyrPheNI 215
 736GACGACAAGCGTTCGTTTCACATTCGCGCCCGTCCAAAGGG 778
 215 sArgArGlnGlySLyAsnTyrValLeuLysIleAlaProProLeuAla 232
 779 AATTC...AACGCAAAAGCCCGATGCGCGCGCGTTCATACCGCAT 825
 232 sPheProSerGlnAspLeuAlaAlaAspThrAlaArgIleAsnGlnVal 248
 826 ACCGAATATGGAATACGCGCTTCGCGACGATATGTTATGTCAA 875
 249 lIleglTrpMetIleAspGluAlaProAspGlnTyrIleTrpIleHisAr 265
 876 CCGCATATAA 885
 265 gArGPheLys 268
 seq_name: sp_bacteria:Q9AP18
 seq_documentation_block:
 ID 09AP18 PRELIMINARY; PRT; 292 AA.
 AC 09AP18:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DE PUTATIVE LAYROYL ACYLTRANSFERASE.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1;
 RX MEDLINE=21142531; PubMed=11208803;
 RA Korotkova N., Lidstrom M.E.;
 RT "Connection between Poly-beta-Hydroxybutyrate Biosynthesis and Growth
 on C1 and C2 Compounds in the Methylophilic Methylobacterium extorquens
 AM1".
 RL J. Bacteriol. 183:1038-1046(2001).
 DR EMBL: AF287907; AAK11538.1; -
 KW transferase; Acyltransferase.
 SO SEQUENCE 292 AA; 32214 MW; 02AD260FA71E7C11 CRC64;

144 ucluleuglyalaargilevalgly.....leuHishislnclglyleg 159
 443 CCATGTACAGCGCCGGAATAATCAAGGATAGCAAAATCATCAGGCG 492
 159 lvaltyltyrargproasnaspasnproleuenastrpilleglphelys 175
 493 GCGACGGTGGCGGCGAAGCAACCGCGCCGCGCATACAGGCGT 542
 176 glyatgilearg.....serasnlysalmetleuasparglyasphe 190
 543 CAACAAATCATCAAGCGCCCTGCGCGCGCGGAGCAACCATCATCTGC 592
 190 uargelymetilleargaleuargasparglyuthtlletprtyalar 207
 593 CGGACCATGCTCTTCCTCCGAGGAGGCGCGCGGTGGGAGATT 642
 207 roasphiaspptyrglyglnln.....asnservalphevalprophe 221
 643 TTGGGCAACCTGCATACACCATGACACTGGCGGCAAAATTGGCACAGT 692
 222 phevalvalgluglnalacysthtrthralaglyserhismetleuarg 238
 693 CAAGGCGTGAACCACTGTTT.....TTCTGCTCGAAGCGCTGCCG 736
 238 gseralaproasnserilevalvalprophetrprolietarghislyra 255
 737 ACGGACAGGCTTCGTGTGCACATCGCGCGCTCCAGGGAATTGAAC 786
 255 spheaserglytyrthvalysilaserproalval...aspheasn 270
 787 GGC.....ACAAGCGCGCATGCGCGGTTCACCGCAATACCGA 830
 271 AspCysglinsnlyslleguthtlralaargleuMetasnlysalvalgl 287
 831 ATATTGCATACGCCGTTTCCGACGAGTATCTGTTATGTACACCGCT 880
 287 uthtlletleuylsalaglnthrglnymettrpneuhsatargp 304
 881 ATAAACG 888
 304 helysthr 306
 seq_name: sp_bacteria:Q9L7U9

seq_documentation_block:
 ID Q9L7U9 PRELIMINARY; PRT; 318 AA.
 AC Q9L7U9;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE LIPID A ACYLTRANSFERASE.
 GN MSBB.
 OS Haemophilus ducreyi.
 OC Bacteria: Proteobacteria: gamma subdivision; Pasteurellaceae;
 OC Haemophilus
 OX NCBI_TaxID=730;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Munson R.S. Jr., Bozue J.A.;
 RT Identification of the msbA gene in Haemophilus ducreyi.
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBD databases.
 DR EMBL: AF219260; AAF33777.1;
 KW Transferase; Acyltransferase.
 SQ SEQUENCE 318 AA; 36799 MW; EACIF82FB88BE37A CRC64;

alignment_scores:
 Quality: 176.00 Length: 263
 Ratio: 1.128 Gaps: 10
 Percent Similarity: 59.316 Percent Identity: 24.715

alignment_block:
 US-09-303-518D-571 x Q9L7U9 ..

Align seg 1/1 to: Q9L7U9 from: 1 to: 318

132 TTCTGCGTTTACCTTTTAAAGAACCGCGCGCGCATCTGCGCAATA 181
 56 serasnlyleuylslysglylysglylysglyntytlyarg. Alaspl 72
 182 TGGCGACGCGGCTTTCAGACCGAC.....ACGAG 213
 72 leasnleucysttyrcyspheproglutrpseralaasplysargalagln 88
 214 ACGTCAAGCGCTTTTGGGGAACGCGCAAAATGCGTTTGGCACTTGC 263
 89 valilecluglnmetpheilleasnvalalaglnthmetleuclglylecl 105
 264 CCCGCGCTTTTCAAAAACCGGAACATCGAACAATGTTCAAAAGCG 313
 105 yglutlilealvalargservalalahlisleuclnlysarpsertuphet 122
 314 TACAGCGCTGGGAACACGTCACACGAGCTTTGGACAGGCGGAGGCTG 363
 122 le...glyleuasnlytyrleuylsglnalalyalysgllyasnille 137
 364 CTGTTCATACGCGCGCATCGGACCTTACGATTGGCGGAGCTATAT 413
 138 lleuileuvalprohistrtrpserilleasplalasergly...illel 153
 414 CAGCGACGACTTCGCTTCACCTGACCGCATGTACAAACCGCGCAAAA 463
 153 eleuHishtrtyrglymetrprometvalsermetlytrsnprohisa 170
 464 TCAACGATAGCAAAATCATGACGCGGCGGAGGCTCGCGCAAGCG 513
 170 snalleuvalasprtleutrpasnalaatrarggluargphelygly 186
 514 AAAACCGCGCCACCGCATACAGGCGGCAAAATCATCAAGCGCT 563
 187 lys...methistrtharglinsnlylleysprophleuasnaspil 202
 564 GCGCGCGCGGAGGACACCATCATCTGCGCGACGCTCTCTTCGCG 613
 202 earglysglyasnmetglytyrphleuproasp.....G 214
 614 AGAAGCGCGC.....GCGCTGTGCGGCGGATTTTTCGCCAAACT 654
 214 lnaspheglygluleuenservaltyralaaspphealatrthglu 230
 655 GCATACACCATGACACTGCGCGCAAAATGCGACAGCGCAAGCGCGAA 704
 231 lysalatrthleuproglyleuasnlysmetalarvalalasnalaagl 247
 705 AACCGTG...TTTTCTGCTCGGAACGCTGCCGACGACAGCGCTCG 751
 247 uvalilleprometpneprolietyrasnalaaglulysglyval...tyrg 263
 752 TGTTCACATCGCGCGCTCCCAAGGGAATTGAACGCGCAAAAGCGCAC 801
 263 lmetglutlleuoproalalle...glupheserglyseretalaagln 278
 802 GATGCGCGCGTGTTCACACCGCAATATGGAATTTGATACGCGGTTTTC 851
 279 seralaaarglunmetasnlysalilleglutyrphevalhishysasntr 295
 852 GACGAGTATCTGTTATGTACAAACCGCTATAAAGC 888
 295 oalaglnlytrvaltrpilleuargleuileuylsthr 307
 seq_name: sp_bacteriap:084013

seq_documentation_block:
 ID 084013 PRELIMINARY; PRT; 455 AA.
 AC 084013;
 DT 01-NOV-1998 (TREMblrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ACYLTFRANSFERASE.
 GN HTFB OR CT010.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D/1W-3/CX;
 RC MEDLINE=9900809; PubMed=9784136;
 RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 RT Chlamydia trachomatis."
 RL Science 283:754-759(1998).
 DR EMBL: AE001275; AAC67600.1;
 KW Transferrase; Acyltransferase; Complete proteome.
 SO SPOUDENCE 455 AA; 52058 MW; 0404B6326C67ACCF CRC64;

alignment_scores:
 Quality: 175.00 Length: 341
 Ratio: 1.054 Gaps: 12
 Percent Similarity: 48.680 Percent Identity: 24.633

alignment_block:
 US-09-303-518D-571 x 084013 ..

Align seg 1/1 to: 084013 from: 1 to: 455

22 CTGTTCCCTTTCGGAACGGCCATGCATCTGTTGACGGCCCTG.. 69
 |||||
 2 LeuphElyArGLeuArGThreLylLysLleuValAspHisLeuVa 18
 70CTCAATGCCCTCTCCCTGCTGCTGCTT 97
 |||||
 18 LlyrGlyLeuGlyLeuGlyValLeuThrLleuArGLeuLeuProArG 35
 98 CCTGTCGACACGCTGGGAAACCGGCTCGGACATCTGGCGTTTACCTT 147
 |||||
 35 eSerLeuArGLeuPheSerLysGlyLeuGlyThAlaLeuPheTybhe 51
 148 TTAAGACAGACCGCGCGCATCGTCCGCAATATG..... 183
 |||||
 52 lIleSerAspValArGlyThrAlaLeuThrAsnLeuAlaLeuAlaPheAr 68
 184CGGACGGCGGTTTGAACCCCGACGACGACGA 214
 |||||
 68 oGluLysPhePheAlaGluArGlyGlnLleAlaArGlnSerValGlnG 95
 215 CGGTCAAGCGCGTTTGGCGAAACGCAAAA..... 246
 |||||
 85 lmeLlelIlelIlePheValGlnLeuAlaThValAspLysPheAlaLys 101
 247TCGGTTTGGAACTTGGCCCC...GC 269
 |||||
 102 HislLeAspLlmeLlIleAlaIleAlaThSerGlnAspAlaProGlu 118
 270 GTTTTCAAAAACCGGACAGACATTCGAACATCTTCAACGCGTACACG 319
 |||||
 118 yPhePhe.....ProGluGluValSerGlnGlnGlnGln..... 130
 320 GCTGGACACAGTGCACAGCGTTTGGACAAAGGCGAGGCGTCTGTTTC 369
 |||||
 131AspHisPhePheSerArGLeuAspArGlnGlnGlnGlnGln 145
 370 ATACGCGGACATCGGACAGTACGATTTGGGGAGACGACATTCAGCA 419
 |||||
 146 PheCysGlyHisGlnAlaAsnTrpGlnLeuProPheLeuTyrlleThr 162
 420 GCACCTTCGTTCCACCTGACCGCCATGTACAAACCGCGCAAAATCAAG 469

162 sArgTyrrProGly.....LeuAlaPheAlaLysProValLysAsnArG 177
 :|||
 470 CGATAGACAAATCATGCAGCGGGGCGGTCGCGCAAGCAAAAC 519
 :|||
 177 rGleAsnGlnLyslIleSerLeuArGLeuSerPheGlnGlyLysle 193
 520 GGGCCACCGCGCATACAAAGGGGTCAAAATCATCAAGCGCTGCGCC 569
 :|||
 194 ValProProGln.....AsnAlaIleAsnGlnAlaLeuAlaLeuHisAr 209
 570 GGGCAGGCAACCATCATCTGCGCCGACCATGTCCTTCGCGAGGAAG 619
 :|||
 209 gGlyGluValValGlyIleValGlyAspGlnValLeuLeuSerSerGut 226
 620 GGGCGGGCGCTGCGCGCATTTTTCGCAAACTGATACCATGACATG 669
 :|||
 226 ySer.....TyrrProLeuPheGlySerGlnAlaPheThrThr 239
 670 CTGGCGGCAAAATTCGACACGTCGCAAGCGCGTGTGTCATCCCGCG 769
 :|||
 240 SerProAlaLeuLeuAlaValLysThrLysThrLysThrValIleAlaValAl 256
 720 CTGCGAAGCGCTGCGCGACGACGTCGCAAGCGCGTGTGTCATCCCGCG 769
 :|||
 256 AlIeTyrrArGlyLysProAsnGlyAsnTyrrLeuValVal..... 268
 770 TCCAGGGAATTCAGACGCAACAAAGCCGACGATGCGCGCTTTCAC 819
 :|||
 269ProSerLysAlaPheHis 274
 820 CGCAATACGAAATTCGATACGC..... 843
 :|||
 275 AlaAsnThrGluLeuSerIleArGlnGlnSerThrGlnGlnLeuMetAspAr 291
 844CGTTT.....CGACGCGATAC 862
 291 gLeuMetArGPhelGluGlnLysGlyLeuThrCysLysProGlnGlnTrpL 308
 863 TGTATTATGTACAAACCGTATATA 885
 :|||
 308 eutTrpLeuHisLysArGTrpLys 315
 seq_name: sp_bacteriap:09PL30
 seq_documentation_block:
 ID 09PL30 PRELIMINARY; PRT: 453 AA.
 AC 09PL30;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE, PUTATIVE.
 GN TC0278.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MOPN / NIGG;
 RC MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwyn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MOPN and Chlamydia
 RT pneumoniae AR39."
 RL Nucleic Acids Res. 28:1397-1406(2000).
 DR EMBL: AE002295; AAF39146.1;
 DR TIGR: TC0278;
 KW Complete proteome.
 SO SEQUENCE 453 AA; 51463 MW; 6221515A0A093FF CRC64;

alignment_scores:

Quality: 162.00 Length: 342
 Ratio: 0.970 Gaps: 13
 Percent Similarity: 48.830 Percent Identity: 23.977

alignment_block:

US-09-303-518D-571 x Q9PL30 ..

Align seg 1/1 to: Q9PL30 from: 1 to: 453

```

22 CTGTTCCCTTGGACCCGATCATCTCTG..... 60
2 Leuphysmetleuargserglylyvalleuvalasphisleuval 18
61 .....ACGCGCGCTCAATGCTCCCTGCTGCGCTT 97
18 Ltyrglyleuglyleuglylleuvalleuargleuvalroarg 35
98 CTTGTCGACACGCTGGAAACGCGTGGACATCGCTTTTACCTT 147
35 erSerleuglnleupheglylysglyleuglythrthrleupheglyval 51
148 TTAAGAGAGACCGCGCGCATCGTCCCATATGCG..... 186
52 lleserisphearglythrallaleuthrasnleuvalaleuvalaphepr 68
187 .....CAGCGCGGTTTGACCCGACACG 211
68 oglytyserpheThrgluargtyrlnlealaleuls...SerValG 84
212 AGACGGTCMAAGCGTTTGGCGAAGCGCAAA..... 246
84 lnclnvalillelthrleuvalaleuvalaspryspheala 100
247 .....TCGGGTTTGACCTTGGCCCC... 267
101 LysHislleasprluillethrleualaserSerGluaspralaprocl 117
268 .CGCTTTTCMAAAACCGGAACATGTCMAAAGCGGTAC 316
117 ugltyrpherpe.....ProgluvalserSerlnglnleu... 130
317 ACGGCTGGGAACGTCGACGCTTGGACAAAGCGGACGCTG 366
131 .....AsnasnpheserArgleuaspGlnleuValalle 144
367 TTCATCAGCGCGACATCGGACATGTTGGCGGACGCTACATCG 416
145 LeupheCysGlnHisGlnAlaAsntrpGluLeuProphelLeuThr 161
417 CGAGGAGCTTCGTTCCACCTGACCGCATATGACCGCGCAAAATCA 466
161 rlysnrlytyrProgly.....LeuAlaPheAlaLysProValLysAsn 176
467 AAGCATGACAAATCATGCGGCGGCGGCGGCGGCAAGGCAAA 516
176 roArgleuAsnArgylsilleSerleuArgLysPheleuLys 192
517 ACGCGCGCGCGCATACAAAGGCGTCAAAATCATCAAGCGCTGCG 566
193 lIleValProProglin...AsnAlaIleAsnGlnAlaLeuArgAlaLeuH 208
567 GCGGCGGCGGACCAATCATCGCGCGGACGCTCTTCCGACG 616
208 slYsglylvalvalIglYlIleValIglYaspIvalleuLeuSerSerg 225
617 AAGCGCGCGCGTGGCGGATTTTTCGCAAACTGATACACATG 666
225 lntYser.....TyrProleupheglyserGlnAlaPheThr 238
667 ACATCGGCGGCAAAATGCGACACGTCMAAGCGCTGAAACCTGTTTT 716

```

```

239 ThrSerProAlaLeuLeuAlaTyrLysThrArgLysProValIleAlaVal 255
717 CTGCTGGAAACGCTGCCGACGACGACAAAGCTTCGTTGTCACATCCGCC 766
255 lAlaIleTyrArgGlnProAsnGlnLysAsnTyrLeuVal..... 268
767 CCGTCGAAGGGGAATTGAACGGCACAAAGCCGACGATCCCGCGGTTC 816
269 .....ProSerLysAlaPhe 273
817 AACCGCAATACGAATATTGATACGC..... 843
274 TyrAlaAsnThrGlnLeuSerIleArgLysThrGlnLeuMetas 290
844 .....CGTGT.....CCGACGAGT 859
290 pLysLeuMetArgPheLeuGlnLysGlylIleAlaLysProGlnLNT 307
860 ATCTGTTTATGTACACCGCTATTA 885
307 rPlenThrPheuHisLysArgTrpLys 315

```

seq_name: sp_bacteriap:Q9ZCL1

seq_documentation_block:

```

ID Q9ZCL1 PRELIMINARY; PRT; 290 AA.
AC Q9ZCL1;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LIPID A BIOSYNTHESIS LAUROYL ACTYLTRANSFERASE (HTRB).
GN RP718.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiaseae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Scheritz-Ponten T., Almark U.C.M., Podowski R.M., Neeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria".
RL Nature 396:133-140(1998).
DR EMBL; AJ235273; CA15149.1; -.
KW Complete proteome.
SQ SEQUENCE 290 AA; 33809 MW; 991FF50AB841D5B3 CRC64;

```

alignment_scores:

Quality: 159.00 Length: 205
 Ratio: 1.214 Gaps: 9
 Percent Similarity: 63.902 Percent Identity: 21.951

alignment_block:

US-09-303-518D-571 x Q9ZCL1 ..

Align seg 1/1 to: Q9ZCL1 from: 1 to: 290

```

289 GACATGAAACAAATGTTCAAGCGGTACAGCGGTGGAAACGTCGACGA 338
95 GluLeuGlnSerArgylleGluille...GlyLysGlnAsnLleLysly 110
339 GCGTTTGGACAGGCGGAAGGCTGCTTCATCAGCGCGCACATCGCA 388
110 sleuGlnuAsp.....GlnProPheLeuLeuPheSerGlnHisPheAla 125
389 GTCAGATTTGGCGGAGCTACATCAGCAGCAGCTTCGTTCCACCTG 438
125 snrPaspIleSerleuLysValleuLysLysTyrTyrPro...LysVal 140

```



```

|||||
255 LeuSerArg1IleSerGlyAlaArg1IleLeuProLeuPhe..... 267
726 AGCGCCGCCGAGGACAA.....GCCTGCTGTTGCACATCCGCCG 769
268 .ProValTyrAspGlyLysThrHisHisLeuThrHisValSerProp 284
770 TCCAAAGGGAATTGAACGGCAACAAAGCCACGATGCCGCCGTTCAC 819
284 roLeuAlaIleArgGlnLysSerAspAlaHisIleAlaArgGlnIleAsn 300
820 GCGAATACCGGATATTGCATACCGCGTTTCCGACGAGTATCTGTAT 869
301 GluValValGluAsnPheValArgProHisProGluGlnTyrThrTrp1 317
870 GTACAAACGCGTATTAACG 888
317 eleuLysLeuLeuLysThr 323
seq_name: sp_bacteria:Q9ZGU3

```

```

seq_documentation_block:
ID Q9ZGU3 PRELIMINARY; PRT; 343 AA.
AC Q9ZGU3;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 10, last annotation update)
DE PUTATIVE ACYLTRANSFERASE.
GN L7029.
OS Escherichia coli O157:H7.
OC Plasmid pO157.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxId=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EDL933;
RC MEDLINE=98391744; PubMed=9722640;
RA Burland V., Shao Y., Perna N.T., Plunkett G., Sofia H.J.,
RT Blatner F.R.;
RT "The complete DNA sequence and analysis of the large virulence plasmid
RT of Escherichia coli O157:H7."
RT Nucleic Acids Res. 26:4196-4204(1998).
DR EMBL; AF074613; AAC70097.1.;
KW Transferase; Acyltransferase; Plasmid.
SQ SEQUENCE 343 AA; 39074 MW; E8869F0D3D30DB6A CRC64;

```

```

alignment_scores:
Quality: 156.50 Length: 273
Ratio: 1.057 Gaps: 10
Percent Similarity: 54.212 Percent Identity: 23.810

```

```

alignment_block:
US-09-303-518D-571 x Q9ZGU3

```

```

Align seg 1/1 to: Q9ZGU3 from: 1 to: 343

```

```

121 CGGCTGGGACATGCGCTTTAACTTTAAAGGAGACCGCGCGCAT 170
:::||||| ||| ||| ||| ||| ||| ||| ||| |||
80 LysLeuGlyMetLeuValGlyArgLeuGlyLysSerAlaArgGlnArgAl 96
171 CGTGGCCATATGCGG.....CAGCGGTTTGAACCCCGACA 208
::: ||| ||| ||| ||| ||| ||| ||| |||
96 aleuIleAsnLeuSerLeuCysPheProGlnTyrSerAspLysGlnLysG 113
209 CCGAGAGGTCAGAGCGCTTTTGGGAAACGCAAAATGCGGTTTGAA 258
::: ||| ||| ||| ||| ||| ||| ||| |||
113 LAsnIleValAspAlaMetPheAlaIleThrAlaSerMetAlaValValLeu 129
259 CTTGGCCCGCGCTTTTCAAAAACCGAGACATGCAAAACATTTTCAA 308
::: ||| ||| ||| ||| ||| ||| ||| |||
130 MetAlaGluLeuAlaLeuSerGlyProAspLysIleSerHisArgIleAr 146

```

```

309 AGCGGTACACGCGCTGGGAACACGTGCAGACGGCTTGGACAAAGCGGAC 358
: ||| ||| ||| ||| ||| ||| ||| |||
146 g...TrpAsnGlyLeuGlnIleValGlyLysMetAlaGlnAsnAsnGln 162
359 GCGTGGCTTTCATCAGCGCCGACATCGGACATGACATTTG..... 399
162 ySValIlePheLeuValProHisAlaIleArgIleValAspIleProAlaMet 178
400 .....GGGAGCGCTTACATCAGCCACGCTTCCGTCCACCT 437
179 LeuMetAlaIleSerGlyArg.....LysMe 187
438 GACCGCCATGTACAGCGCCGCAAAATGCAAGAGATGACAAATCATGC 487
187 tAlaAlaMetPheHisAsnGlnArgAsnProValValAspTyrValTrpA 204
488 AGCGGGCGAGGTGCGCGCAAAAGCAAAACCGCCGACCGCATACAA 537
204 sSerValArgArgArgPheGlyGlyLysLeuHisAlaArgAsn...Asp 219
538 GGGGTCAAAACATTCATCAGGCGCTGCGCGGCGGAGGCAACCATCAT 587
220 GlyIleAlaSerPheValArgSerValArgGlnGlyTyrTrpGlyTyrTy 236
588 CTTGCCC.....GACCAAGTCCCTTCTCCGACGAGAGCGGCGGTGT 631
236 leuProAspGlnAspHisGlyProGluPheSerGlu.....P 249
632 GGGCGATTTTTCGGCAAAACGTCATACCATGACATGCGCGGCAAA 681
249 heAlaAspPhePheAlaThrTyrLysAlaThrLeuProValIleGlyArg 265
682 TTGGCACACGTCAAAAGCGGTGAA.....ACCTGTTTTCGTGGCA 725
266 LeuSerArgIleSerGlyAlaArgIleIleProLeuPhe..... 278
726 AGCGCTCCCGGACGAGCAAA.....GCCTGCTGTTGCACATCCGCCG 769
279 .ProValTyrAspGlyLysThrHisHisLeuThrHisValSerProp 295
770 TCCAAAGGGAATTGAACGGCAACAAAGCCACGATGCCCGCGTTTCAAC 819
295 roLeuAlaIleArgGlnLysSerAspAlaHisIleAlaArgGlnIleAsn 311
820 GCGAATACCGGATATTGCATACCGCGTTTCCGACGAGTATCTGTAT 869
312 GluValValGluAsnPheValArgProHisProGluGlnTyrThrTrp1 328
870 GTACAAACGCGTATTAACG 888
328 eleuLysLeuLeuLysThr 334

```

```

seq_name: sp_bacteria:Q9AKA6

```

```

seq_documentation_block:
ID Q9AKA6 PRELIMINARY; PRT; 290 AA.
AC Q9AKA6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE HTRB. PROTEIN.
GN HTRB.
OS Rickettsia typhi.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiinae; Rickettsia.
OX NCBI_TaxId=785;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINNINGTON.
RC MEDLINE=21219194; PubMed=11319266;
RA Andersson J.O., Andersson S.G.E.;
RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes."

```


RL Mol. Biol. Evol. 18:829-839(2001).
 DR EMBL: AJ293328; CAC3762.1; .. 2E9B2D3193B70EDD CRC64;
 SQ SEQUENCE 290 AA; 33814 MW; 2E9B2D3193B70EDD CRC64;

alignment_scores:
 Quality: 155.50 Length: 195
 Ratio: 1.275 Gaps: 7
 Percent Similarity: 62.564 Percent Identity: 22.051

alignment_block:
 US-09-303-518D-571 x Q9AKA6 ..

Align seg 1/1 to: Q9AKA6 from: 1 to: 290

```

319 GCGTGGGACACGTGACGAGGCTTTGGACAGGCGGACAGGCTGCTGT 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
104 GLYSGIAsnIleGlnIleuGlnuAsp.....GlnProPheLeu 118
    : : : : : : : : : : : : : : : : : : : : : : : : : :
369 CATCGCGCGGACATCGGACGTACGATTGGCGGACGCTACATCAGCC 418
    : : : : : : : : : : : : : : : : : : : : : : : : : :
118 uPheSerGlyHisPheAlaAsnTTPAsPIleSerIleuValIleuYSL 135
    : : : : : : : : : : : : : : : : : : : : : : : : : :
419 AGCAGCTTCGCTCCACCTGACCGCCATGACAAAGCCGCAAAATCAA 468
    : : ||| : : : : : : : : : : : : : : : : : : : :
135 YSTYTYTPro...LysValAlaValIleTYrArgLysAlaAsnAspPro 150
    : : ||| : : : : : : : : : : : : : : : : : : : :
469 GCGATAGACAAATATCATCGAGCGGCGGAGGCGCGGCAAGGCAAAAC 518
    : : : : : : : : : : : : : : : : : : : : : : : : : :
151 TYrValAsnIleuValAsnIleuSerIleuArgIleuAspLysIleuArg 167
    : : : : : : : : : : : : : : : : : : : : : : : : : :
519 CCGCGCCACCGGACATACAGGGGTCAAAACAAATCATCAGGCGCTGCGCG 568
    ||| ||| : : : : : : : : : : : : : : : : : : : :
167 uIleProLysGlySerGlnGlyIleArgAlaIleuValSerAlaIleuYSG 184
    : : : : : : : : : : : : : : : : : : : : : : : : : :
569 CCGCGGAGGCAACCATCTCTGCGCGGACGCTGCTTCCGCGAGGAA 618
    ||| ||| : : : : : : : : : : : : : : : : : : : :
184 IuGlyIleuSerIleuValIleuValAspGln.....LysMet 196
    : : : : : : : : : : : : : : : : : : : : : : : : : :
619 GCGCGGCGGCTGAGGCGGATTTTTCGCAAAACCTGATACAGCATGAC 668
    : : : : : : : : : : : : : : : : : : : : : : : : : :
197 AsnAspGlyIleGlnValProPheLeuGlyArgProSerMetThAlaSe 213
    : : : : : : : : : : : : : : : : : : : : : : : : : :
669 ACTGGCGCAAAATTTGGACAGCTCAAGGCGTGAAGAACCTGTTTCT 718
    ||| ||| : : : : : : : : : : : : : : : : : : : :
213 IAlaIleAlaLysIleAla.....LeuGlnTYrLysTYrProIleLeP 228
    : : : : : : : : : : : : : : : : : : : : : : : : : :
719 GCTGGCAAGCGCTG...CCGAGGCAAGGCGTCTGTTGCACATCGCG 765
    : : : : : : : : : : : : : : : : : : : : : : : : : :
228 rocysGlnIleIleArgThLysGlySerTYrPheLysValIleValHis 244
    : : : : : : : : : : : : : : : : : : : : : : : : : :
766 CCC...GTCCAGGCGAATTTGAAGGCAACAAAGCCGACAGTCC..... 807
    ||| ||| : : : : : : : : : : : : : : : : : : : :
245 ProGlnIleuLysPheGlnIleuGlyAspAsnLysValAspCysTYrAs 261
    : : : : : : : : : : : : : : : : : : : : : : : : : :
808 .....GCCGTGTTCAAGCGGCAATACCGAATATGTGATACGCGCTTTC 850
    : : : : : : : : : : : : : : : : : : : : : : : : : :
261 nIleMetPheTYrIleAsnGlnIleLeuGlyIuTYrValLysGlnAsp 278
    : : : : : : : : : : : : : : : : : : : : : : : : : :
851 CGAGCGAGTATGTTATGTATACACCGCTATAA 885
    ||| ||| : : : : : : : : : : : : : : : : : : : :
278 roAlaGlnTYrPheTYrPheHisAsnArgTYrLys 289
    : : : : : : : : : : : : : : : : : : : : : : : : : :

```

seq_name: sp_bacteriap:Q9CME5

seq_documentation_block:

ID Q9CME5 PRELIMINARY: PRT: 313 AA.
 AC Q9CME5:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHEtical PROTEIN PM0885.
 GN PM0885.

OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Pasteurella.
 OX NCBI_Taxid=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L., Paustian M.L., Whitlam T.S., Kapur V.,
 RT "Complete genomic sequence of Pasteurella multocida Pm70."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL: AF006127; AK02969.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 313 AA; 35956 MW; 23456DB87C45E535 CRC64;

alignment_scores:
 Quality: 154.00 Length: 287
 Ratio: 0.975 Gaps: 10
 Percent Similarity: 55.052 Percent Identity: 25.087

alignment_block:
 US-09-303-518D-571 x Q9CME5 ..

Align seg 1/1 to: Q9CME5 from: 1 to: 313

```

64 GCGCTGCTCAAAATGCTCTCCCTGCTGCGCTTCTCTGTCACAGCT 113
    ||| ||| : : : : : : : : : : : : : : : : : : : :
26 AlalaTYrArgIleuIleValCysLeuProTYrProIleuHisLysLe 42
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114 GGGAAACGGGCTCGACATCTGCGGCTTTTAAAGGAGACCGCG 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
42 uGlyAsnGlyLeuGlyTYrLeu.....LeuAlaLysSerThValG 56
    : : : : : : : : : : : : : : : : : : : : : : : : : :
164 CCGGCAATGCTCGCAATATGCGGCGGCGGCTTTGAC..... 201
    : : : : : : : : : : : : : : : : : : : : : : : : : :
56 IyGlnArgArgAlaAsnIleAlaArgArgAsnLeuGlnLeuCysPhePro 72
    : : : : : : : : : : : : : : : : : : : : : : : : : :
202 .....CCGACAGCGAGAG...GTCAAAACGCTTTTTCGCGA 236
    : : : : : : : : : : : : : : : : : : : : : : : : : :
73 HisTYrSerAlaAspGlnIleGlnIleuThLeuHisAlaHisLeuYsAl 89
    : : : : : : : : : : : : : : : : : : : : : : : : : :
237 AACGGCAAAATGCGGTTTGGAACTTGGCCCGGCTTTTCAAAAACCG 286
    ||| ||| : : : : : : : : : : : : : : : : : : : :
89 aThrGlyMetAlaIleIleGlnIleuThGlyMetAlaTYrPheTYrSer 106
    : : : : : : : : : : : : : : : : : : : : : : : : : :
287 AAGACATGCAAAATGCTCAAAAGCGGACAGCTGGAACAGTGCAG 336
    ||| ||| : : : : : : : : : : : : : : : : : : : :
106 IArgIleLysLysTYrSerLys...IleGlnGlyLeuGlyTYrLeuLys 121
    : : : : : : : : : : : : : : : : : : : : : : : : : :
337 CAGGCTTGGACAAAGGCGAAGGCTGCTCTTCAATCAAGCGGCAATCG 386
    : : : : : : : : : : : : : : : : : : : : : : : : : :
122 GlnTYr.....LysGlnAspGlyValIleLeuValGlyValHisPheLe 136
    : : : : : : : : : : : : : : : : : : : : : : : : : :
387 CAGCTACGATTTGGCGGAGCTACATCAGCCAGCAGCTTCCGTCAC 436
    : : : : : : : : : : : : : : : : : : : : : : : : : :
136 uThrLeuIleuGlnGlyAlaArgIleValGlyLeuHisIleProLys... 151
    : : : : : : : : : : : : : : : : : : : : : : : : : :
437 TGACCGCATGACAGCGCGGCAAAATCAAGAGATAGCAAAATCATG 486
    : : : : : : : : : : : : : : : : : : : : : : : : : :
152 ..IleGlyValTYrArgProAsnAspAsnProLeuMetAspTYrLeuGln 167
    : : : : : : : : : : : : : : : : : : : : : : : : : :
487 CAGCGGCGAGGCTGCGGCAAGGCAAAACCGGCGGCGGCAACATCA 536
    ||| ||| : : : : : : : : : : : : : : : : : : : :
168 ValGlnGlyArgLeuArgSerAsnLysAspLeuAspArg.....Ly 182
    : : : : : : : : : : : : : : : : : : : : : : : : : :
537 AGGGCTCAAAATCATCAAGCGCTGCGGCGGCGGCGGCAACATCA 586
    : : : : : : : : : : : : : : : : : : : : : : : : : :
182 sasPleuArgGlyMetIleArgAlaLeuArgGlnIleuIleIleTYrPT 199
    : : : : : : : : : : : : : : : : : : : : : : : : : :
587 TCCTGGCGGACAGCTCTTCCGAGAGAGGCGGCGGCTGTCGCG 636
    ||| ||| : : : : : : : : : : : : : : : : : : : :
199 yAlaIleProAspHis.....AspTYrGlyArgLysAlaValAlaPheVal 213
    : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

637 GATTTTTCGGCAACCTGCATACACCATGACATGCGGCGCA.....AA 680
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 ProPheheValProAspAlaIaThrThrThrGlySerTyrTle 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 ATTGGCACAGCTCAAGCGCTGAAACCCCTTTTTCGTGCGCAAGCC 730-
|||||:|||||:|||||:|||||:|||||:|||||:
230 uLeuYSerSerProThrSerTyrValIleProPheAlaProLeuArg 247
731 TGCCGACGACGACGCTGCTTGGCAATCCG...CCCGTCCAGGG 777
|||||:|||||:|||||:|||||:|||||:|||||:
247 snserAspGlySerGlyTyrThrValIleSerProProValAspPhe 263
778 GAATTGAACGGCAACAAAGCCGATGCGCGCTTTTCAACGCAATAC 827
|||||:|||||:|||||:|||||:|||||:|||||:
264 SerAspIleThrAspGluThrArgIleAlaThrArgMetAsnGlnVal 280
828 CGAATATGATACGCGCTTTTCGACGCGATATCTGTATATACACC 877
|||||:|||||:|||||:|||||:|||||:|||||:
280 IGluThrGluIleMetLeuSerGlyValGluGlnTyrMetLeuHisArg 297
878 GCTATAAACG 888
|||||:|||||
297 rghelysthr 300

```

seq_name: sp_bacteria:Q9AKF0

```

seq_documentation_block:
ID Q9AKF0 PRELIMINARY; PRT; 290 AA.
AC Q9AKF0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HTRB PROTEIN.
GN HTRB.
OS Rickettsia rickettsii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxId=783;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=84-21C;
RX MEDLINE=21219194; PubMed=11319266;
RA Andersson J.O., Andersson S.G.E.;
RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes.";
RL Mol. Biol. Evol. 18:829-839(2001).
DR EMBL; AJ293329; CAC33715.1;
SO SEQUENCE 290 AA; 33550 MW; 9FC5D73E5CBE89A CRC64;

```

alignment_scores:

Quality:	149.50	Length:	198
Ratio:	1.225	Gaps:	7
Percent Similarity:	61.616	Percent Identity:	22.222

alignment_block:

US-09-303-518D-571 x Q9AKF0

Align seg 1/1 to: Q9AKF0 from: 1 to: 290

```

310 GCGGTACAGCGCTGGAAACAGCTGACAGCTTTGGCAAGGCGAGG 359
|||||:|||||:|||||:|||||:|||||:|||||:
101 AlaIleIleGlyIleGluAsnIleLysLeu.....GluGlyGlnPrl 115
360 GCTGTGTCATACGCGCGACATGCGACATGATTTGGCGGAGCGCT 409
|||||:|||||:|||||:|||||:|||||:|||||:
115 orheuleuPheSerGlyHisPheAlaAsnTrpAspIleSerIleLysI 132
410 ACATTCAGCCAGAGCTTCGTTCCATCGACGCGCATGTACAGCGCGG 459
|||||:|||||:|||||:|||||:|||||:|||||:
132 IeLeuHisLysPheTyrPro...LysValAlaValIleTyrArgLysAla 147
460 AAATCAAGCATGACAAATATCATGACGCGGAGGCTGCGGCGCA 509

```

```

148 AsnAsnProTyrValAlaSnLysLeuValAsnGluSerArgAlaGlyAspTyr 164
|||||:|||||:|||||:|||||:|||||:|||||:
510 AGGCAAAACCGCGCCACCGCATACAGGGGTCAAAACAAATCATCAG 559
|||||:|||||:|||||:|||||:|||||:|||||:
164 sLeuArgLeuIleProLysGlyProGluLysSerArgAlaLeuValArg 181
560 CCGTGGCGCGGCGGAGGACACATCATCTGCCGACCGCTCCCTTCT 609
|||||:|||||:|||||:|||||:|||||:|||||:
181 laIleLysGluSerGluAlaIleValMetLeuValAspLln..... 194
610 CCGCAGAGAGCGCGCGCGCTGTGGGCGGATTTTTCGCGCAACCTGCAT 659
|||||:|||||:|||||:|||||:|||||:|||||:
195 ...LysMetAsnAspLysIleGluValProPheLeuGlyHisProAlaMe 210
660 CACCATGACACTGCGCGCAAAATGCGACAGCTCAAAAGCGTGAAACCC 709
|||||:|||||:|||||:|||||:|||||:|||||:
210 tThrAlaSerAlaIleAlaLysIleAla.....LeuGlnTyrLysTyrP 225
710 TGTTTTTCGTGCTGCGACGCGCTG...CCGACGCAAGAGCTTCGTTG 756
|||||:|||||:|||||:|||||:|||||:|||||:
225 roIleIleProCysGlnIleIleArgThrLysGlySerTyrPheLysVal 241
757 CACATCCGCGCC...GTCCAGGCGAATGGAAGCGCAACAAAGCCAGCA 803
|||||:|||||:|||||:|||||:|||||:|||||:
242 ILeValHisProGlnLeuLysLeuGlnThrGlyAspAsnLysAlaAs 258
804 TGCC.....GCCGTTCACCGCAATACCAATATGATGATAC 841
|||||:|||||:|||||:|||||:|||||:|||||:
258 pCysTyrAsnIleMetLeuAsnIleAsnGlnMetLeuGlyLutPValI 275
842 GCGCTTTCCGACGACATATCTGTTATGTACAAACCGTATATAA 885
|||||:|||||:|||||:|||||:|||||:|||||:
275 ySgIAsnProAlaGlnTrpPheThrPheHisAsnArgTyrLys 289

```

seq_name: sp_bacteria:P94909

```

seq_documentation_block:
ID P94909 PRELIMINARY; PRT; 529 AA.
AC P94909;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 58.7 KDA PROTEIN (FRAGMENT).
OS Microbacterium ammoniaphilum.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococccineae; Microbacteriaceae; Microbacterium.
OX NCBI_TaxId=42460;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15354;
RX MEDLINE=96527250; PubMed=8654988;
RA Striebel H.M., Seebert S., Jarsch M., Kessler C.;
RT "Cloning and characterization of the Mami restriction-modification
RT system from Microbacterium ammoniaphilum in Escherichia coli.";
RL gene 172:41-46(1996).
DR EMBL; X79027; CAA55649.1;
KW Hypothetical protein.
FT NON_TER
SO SEQUENCE 529 AA; 58716 MW; 5DC1216237DF772D CRC64;

```

alignment_scores:

Quality:	146.00	Length:	320
Ratio: <th>1.090</th> <td>Gaps:</td> <td>19</td>	1.090	Gaps:	19
Percent Similarity:	41.875	Percent Identity:	27.188

alignment_block:

US-09-303-518D-571 x P94909

Align seg 1/1 to: P94909 from: 1 to: 529

```

30 CCCTTTCGCAACCGCATGACATCTCTGTGACGCGCTGTCAATATGCC 79

```

```

311 GlycylalalalargleuaspargalaglGlyValGlyAlaArgProAl 327
815 TCACCCGCA 824
      ::::|:::
327 ArgProArg 330

seq.name: sp_invertebrate:Q9YU43

seq_documentation_block:
ID      Q9YU43      PRELIMINARY;      PRT;      954 AA.
AC      Q9YU43;
DE      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      CG11274 PROTEIN.
GN      CG11274.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC      Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BERKELEY;
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA      Abil J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A.A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Beres P.V., Berman B.P., Bhandari D., Bolintsov S.,
RA      Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA      Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de la Pallos B., Delcher A., Deng Z., Dwyer A.D., Dew I., Dietz S.M.,
RA      Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA      Folsler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA      Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA      Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA      Jatalin D., Kalush F., Kaplen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA      Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA      Lasero P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Meltai B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Pollard J., Pui V., Reese M.G.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
RA      Relhart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA      Spletter E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA      Svirskas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA      Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA      Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA      Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA      Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT      "The genome sequence of Drosophila melanogaster.";
RL      Science 287:2185-2195(2000).
DR      EMBL, AE003538; AAF9848.1; -
DR      FlyBase; FBgn0036340; CG11274.
DR      InterPro; IPR002483; PWT.
DR      Pfam; PF01480; PWT. 1.
DR      SMART; SM00311; PWT. 1.
DR      PROSITE; PS00055; RIBOSOMAL_S12; UNKNOWN_1.
SQ      SEQUENCE 954 AA; 107592 MW; 5722165979110A107 CRC64;

```

alignment_scores:
 Quality: 145.00 Length: 329
 Ratio: 1.074 Gaps: 17
 Percent Similarity: 41.033 Percent Identity: 25.228

alignment_block:

US-09-303-518D-571 x Q9VU43 ..

Align seg 1/1 to: Q9VU43 from: 1 to: 954

```

150 AAAGAACACCGCGCGCATGTCGCAATATCGCGAGCGG.....194
    :::::  :::::  :::::  :::::  :::::  :::::  :::::
347 ArgGlnAsnAraArgGly.....ArgGlnPheGlyGlyGlyGly1 361
195 .....TTGAACCCGACAGCAGAC.....GGTC 219
    :::::  :::::  :::::  :::::  :::::  :::::
361 YArgArgGlyLysArgArgPheSerProArgGlyArgSerProMetArg 378
220 AAAGCCGTTTTCGCGAAACGCAAAATGCGGTTT.....254
    :::::  :::::  :::::  :::::  :::::  :::::
378 LAspAsnGlyGlyLysAsnGlyLysAsnArgPheGlnArgSerAsnSer 394
255 .....GGAACCTGCCCCGCGCTTTTCA 277
395 ArgArgArgSerArgSerArgArgLeuSerArgSerProMetArgTyrSe 411
278 AAAAACCGAGACATCGAAACATGTTCAAGCGGTACACGCGTGGGAA 327
    :::::  :::::  :::::  :::::  :::::  :::::
411 rArgSerProArgArgPheAsnAsnArgArgArgSerProMetArgTyrSe 428
328 CAGCTGCACGAGGCTTGGACAAAGGCGGAGGCTGTTCATCACGCC 377
    :::::  :::::  :::::  :::::  :::::  :::::
428 heArgGlyGlyLysArgGlyGlyGlyLysArg.....GlyGly 439
378 GCACATGCGCAGCTA.....CGATT 397
    :::::  :::::  :::::  :::::  :::::  :::::
440 GlyHisArgGlnMetTrpGlnHisArgGlySerProAsnPheArgGly 456
398 TGGCGGAGCGCTACATCAGCCAGCGCTCCGTCACCT.....437
    :::::  :::::  :::::  :::::  :::::  :::::
456 yGlyLysArgGlnHisTrpGlnAlaArgHisSerProGlyGlyGlyLys 473
438 .....GACCGCCATGTACAAAGCCGCC 458
    :::::  :::::  :::::  :::::  :::::  :::::
473 rGlyGlyGlyGlyGlyMetGlyArgPheAsnArg.....484
459 GAAATCAAGCGATAGCAAAATCATGACGCGGAGGCGGCGCGCA 508
    :::::  :::::  :::::  :::::  :::::  :::::
485 ..ArgGlnSerProGlnGlnAsn.....491
509 AAGCGAAACCGCGCCACCGCATACAGGGGTCAAACAAT.....551
    :::::  :::::  :::::  :::::  :::::  :::::
492 ArgTyrAsnArgPheHisArgGlnSerProMetGlnGlnAsnGlnProp 508
552 .....CATCAAGCCCTGCGCGCGGCGAGCGCAAC 581
    :::::  :::::  :::::  :::::  :::::  :::::
508 heArgLysGlnPheSerProHisGlnGlyProGlyArgArgPheSerSer 524
582 CATCATCTCGCCGA.....CCACGTCCTTCCTCCGAGAGAG 619
    :::::  :::::  :::::  :::::  :::::  :::::
525 ProGlnGlnArgArgAsnSerArgAsnArgArgProAsnSerArgGluArg 541
620 GCGGCGGCGTGTGGCGGATTTTTCGCAAACTGCATACACCATGACA 669
    :::::  :::::  :::::  :::::  :::::  :::::
541 gArgSerSerProGlyGly.....547
670 CTGGCGCAAAATTTGGACACAGCTCAAGGCGTGAACCTGTTTCTG 719
    :::::  :::::  :::::  :::::  :::::  :::::
548 ..GlyGly.....GlyHisMetAsnArgTrpAsnPro.....558
720 CTGGCAACGCGCTGCC.....734
559 .....ProProAlaArgAsnArgArgSerSerSerGlySerSerAlaGly 573
  
```

```

735 ....CGACGACAAAGGCTTCGTTCACATCCGCCC.....767
    :::::  :::::  :::::  :::::  :::::  :::::
573 yGlyArgGlnGlnArgGlnArgSerProProGlyLysArgArgS 590
768 .....CGTCCAGGAGGAAATGAAACGCAACCAAGCCAGATGC 806
    :::::  :::::  :::::  :::::  :::::  :::::
590 erSerHisSerArgSerArgSerLeuSerArgAsnArgSer...ArgSer 605
606 ArgSerSerSerGlnHisSerArgLysArgGluSerProGlyGlyArgSe 622
857 AGTATCTGTTATGTACAAACCGGTATAAACCCGCTA 893
    :::::  :::::  :::::  :::::  :::::  :::::
622 rSerValGlyTyrAlaGlyProAlaValAsnThrIle 634
seq_name: sp_bacteriap:Q92GN2
  
```

```

seq_documentation_block:
ID Q92GN2 PRELIMINARY; PRT; 290 AA.
AC Q92GN2;
DT 01-DEC-2001 (TREMBLrel, 19, Created)
DR 01-DEC-2001 (TREMBLrel, 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel, 19, Last annotation update)
DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.
GN HTRB OR RC1091.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiase; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samsom D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
DR EMBL; AE008660; AAL03629.1;
KW Science 293:2093-2098(2001).
KW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 290 AA; 33522 MW; 1C0FD05318332FD0 CRC64;
  
```

alignment_scores:
 Quality: 144.50 Length: 198
 Ratio: 1.194 Gaps: 7
 Percent Similarity: 61.111 Percent Identity: 21.717

alignment_block:

US-09-303-518D-571 x Q92GN2 ..

Align seg 1/1 to: Q92GN2 from: 1 to: 290

```

310 GCGGTACAGCGCTGGGACACGTCGACAGCGTTTGACAAAGCGGAAG 359
    :::::  :::::  :::::  :::::  :::::  :::::
101 AlaIleIleGlyIleGlnAsnIleLysLysLeu.....GlnGlyGlnProp 115
360 GCTGCTGTTTCATCAGCGCGCACATCGGCGAGCTAGATTTGGCGGAGCCT 409
    :::::  :::::  :::::  :::::  :::::  :::::
115 orPheLeuLeuPheGlyGlyHisPheAlaAsnTrpPheSerIleLysI 132
410 ACATCAGCAGCAGCTCCGTTCCACCTCAGCCGCGCATGACAAAGCGCGC 459
    :::::  :::::  :::::  :::::  :::::  :::::
132 LeuHisLysPheTyrPro...LysValAlaValIleLysTrpGlySAla 147
460 AAATTCAAAGCGATAGACAAATATGACGCGCGGCGAGGTCGCGGCA 509
    :::::  :::::  :::::  :::::  :::::  :::::
148 AsnAsnProTyrValAsnLysLeuValAsnGlnSerArgAlaGlyAspLys 164
510 AGCGAAACCGCGCCCGCATACAAAGGCTCAAAACATCATCAAG 559
    :::::  :::::  :::::  :::::  :::::  :::::
  
```

```

164 sLeuArTgIeuIleProLysGlyProGluGlySerArgAlaLeuValArgA 181
560 CCGTCGGCGGGGCGGAGGACCATCATCTGCGCCGACGAGCTGCTCT 609
181 lIleLysGluSerGluSerIleValMetLeuValAspGln..... 194
610 CCGCAGGAGGCGGGCGGTGTGGCGGATTTTTCGGCAAACTGCATG 659
195 ...LysMetAsnAspGlyIleGluValProPheLeuGlyHisProAlaMe 210
660 CACCATGACACTGGCGGCAAAATTTGGCACAGCTCAAGCGCGTGAACCC 709
210 tThrAlaSerAlaIleAlaLysIleAla.....LeuGlnTyrLysTyrP 225
710 TGTATTTCGCTGCGACGCGCTG...CCGACGAGCAAGGCTGCTGTG 756
225 rolleIleProCysGlnIleIleArgThrLysGlySerTyrPheLysVal 241
757 CACATCCGCCCC...GTCCAGGAGGAATTGAAGGCAACAAAGCCACGA 803
242 lIleValHisProGlnLeuLysPheGlnThrGlyAspAsnLysAlaAs 258
804 TGGC.....GCCGTGTCAACCGCAATACCGAATTTGGATAC 841
258 pCysTyrAsnIleMetLeuAsnIleAsnGlnMetLeuGlyGluTyrVal 275
842 GCCGTTTCCGACGAGTATCTGTATGATACACCGCATATAA 885
275 yGlnAsnProAlaGlnIleThrPheThrPheHisAsnArgTyrLys 289

```

seq_name: sp_bacteria:Q9AKK5

```

seq_documentation_block:
ID Q9AKK5 PRELIMINARY; PRT; 290 AA.
AC Q9AKK5:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HTRB PROETIN.
GN HTRB.
OS Rickettsia montana.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxId=33991;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21219194; PubMed=11319266;
RA "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes.";
RT Mol. Biol. Evol. 18:829-839(2001).
RL EMBL; AJ293330; CAC33650.1;
DR SEQUENCE 290 AA; 33635 MW; 9599E3C0C3C076P9 CRC64;

```

alignment_scores:

Quality: 142.50 Length: 195
Ratio: 1.197 Gaps: 7
Percent Similarity: 61.026 Percent Identity: 21.538

alignment_block:

US-09-303-518D-571 x Q9AKK5 ..

Align seg 1/1 to: Q9AKK5 from: 1 to: 290

```

319 GCGTGGACACAGTCGACGAGCTTTGACAGGCGGAAGGCGCTGCT 368
||| |||:|||||:|||||:|||||:|||||:|||||:|||||:
104 GlyIleGluAsnIleLysLysLeu.....GluGlyGlnProPheLeu 118
369 CATCACGCGGACATCGGACGCTACGATTTGGGCGAGCATACAGCC 418
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
118 upSerGlyHisPheAlaAsnTyrPsrPheSerLeuLysLeuHisL 135
419 AGCAGCTTCGCTTCCACCTGACCGGCGCATGTACAGCGCGCAAAATCAA 468

```

```

135 ySPheTyrPro...LysValAlaValIleTyrArgLysAlaAsnPro 150
:: ||| :|||||:|||||:|||||:|||||:|||||:
469 GCGTATACAAATTCATGCGGCGGAGGTCGCGGCAAGGCAAAAC 518
151 TyrValAsnLysLeuValAsnGluSerArgAlaGlyAspLysLeuArg 167
519 CCGGCGCCACCGCATACAGGCGGTCAACAAATCATCAAGGCGCTGCGG 568
||| ||| :|||||:|||||:|||||:|||||:|||||:
167 uIleProLysGlyProGluGlySerArgAlaLeuValArgAlaIleLysG 184
569 CCGGCGGACGACCATCATCTGCGCGACAGCTCCCTTCGCGCAGGA 618
184 lSerGluProIleValMetLeuValAspGln.....LysMet 196
619 GCGGCGCGCGTGGCGGCGGATTTTTCGGCAAACTGCATACACCTGAC 668
||| |||:|||||:|||||:|||||:|||||:|||||:
197 AsnAspGlyIleGluValProPheLeuGlyHisProAlaMetThrLase 213
669 ACTGCGCGCAAAATTTGGCACAGCTCAAGGCGTGAAGCCCTGTTTCT 718
: |||:|||||:|||||:|||||:|||||:|||||:
213 rAlaIleAlaLysIleAla.....LeuGlnTyrLysTyrProIleIleP 228
719 GCTGCGACCGCTG...CCGACGAGCAAGGCTTCGTTGCGACATCCG 765
|||:|||||:|||||:|||||:|||||:|||||:
228 rOCysGlnIleIleArgThrLysGlySerTyrPheLysValIleValHis 244
766 CCC...GTCCAGGAGGAATTGAAGGCAACAAAGCCACGATGCC..... 807
|||:|||||:|||||:|||||:|||||:|||||:
245 ProGlnLeuLysPheGlnGlnThrGlyAspAsnLysAlaAspCysTyrAs 261
808 .....GCCGTGTCAACCGCAATACCGAATTTGGATACGCCGTTTTC 850
|||:|||||:|||||:|||||:|||||:|||||:
261 nIleMetLeuAsnIleAsnGlnMetLeuGlyGluTyrValLysGlnAsn 278
851 CGACGCGATATCTGTATGATACACCGCATATAA 885
|||:|||||:|||||:|||||:|||||:|||||:
278 rOserGlnTyrPheThrPheHisAsnArgTyrLys 289

```

seq_name: sp_human:Q16824

```

seq_documentation_block:
ID Q16824 PRELIMINARY; PRT; 797 AA.
AC Q16824:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROFLAGGRIN (FRAGMENT).
GN FLAG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91064347; PubMed=2248957;
RA Gan S.Q., McBride O.W., Idler W.W., Markova N., Steinert P.M.;
RT "Organization, structure, and polymorphisms of the human profilaggrin
RT gene [published erratum appears in Biochemistry 1991 Jun
RT 11;30(23):5814].";
RL Biochemistry 29:9432-9440(1990).
DR EMBL; M60502; AAA63248.1;
DR InterPro; IPR003303; FilaGrIn.
DR PRINTS; PR00487; FilaGrIN.
FT NON_TER 1
SEQUENCE 797 AA; 85176 MW; 60B6184763BDA86B CRC64;

```

alignment_scores:

Quality: 142.00 Length: 222
Ratio: 1.193 Gaps: 10
Percent Similarity: 53.604 Percent Identity: 24.775

alignment_block:
us-09-303-518d-571 x Q16824 ..

Align seg 1/1 to: Q16824 from: 1 to: 797

```

162 CGCCGCGATCGTCCCAATATGCGGCGGCGTTTGAACCCGACACGC 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 ArgThtSerArgArgIn.....GlySerValSerGlnAspSe 279
212 AGACGGTCAAGCGTTTGGGAAAGCGCAAAATCGCGTTTGAACCT 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
279 rAspSerGlnGlnHis...SerGlnAspSerGlnArgArgSerGlySe 295
262 GCCCGCGCGTTTTCAAAAAACCGGAAACATCGCAAAATGTTCAAGC 311
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
295 lAspArg.....AsnHisArgGlySerAlaGlnGln 305
312 GGTACACGCGTGGGAAACGCT.....GCAGC 337
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
306 GlnSerArgAspGlySerArgHisProArgSerHisGlnAspArgAl 322
338 AGCGTTTGA.....CAAGCGGAAAGCGCTGCTGATCAGC 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
332 acLynHisLysPserAlaGlnSerArgGlnSerGlyThrHisHisA 339
376 CGGCACATCGGCACTACGATTTGGCGGCGAGCTACATCAGCCAGACT 425
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339 lAGlnAspSerSerGlyGlnAlaAlaSerSerHisGlnGlnAlaArg 355
426 TCCGTTCCACCTGACCGCATGTACAAAGCCCGCAAAATCAAGCGATAG 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
336 SerSerAlaGlyGlnArgHisGlySerHisTyrGlnGlnSerAlaAsp 372
476 ACAAAATCATCGACGC.....GGCGAGGTCGCGCGCAAGCGCAA 516
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
372 rSerArgHisSerGlyLeuGlnHisGlyGlnAlaSerSerAlaValArg 389
517 ACCGCGCGCCCGCATACAGGGCTCA.....ACAAATCATCAAGGC 560
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
389 sPserGlyHisArgGlySerSerGlySerGlnAlaSerAspAsnGlnGly 405
561 COTGCG.....CGCGGCGCAGC 577
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 HisSerGlnAspSerAspThrGlnSerValSerAlaHisArgGlnAlaG 422
578 CAACCATCATCTCGCCCA.....CCAGTCCTCTTCCCGCAGGAA 618
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
422 YArgHisHisGlnSerHisGlnGlnSerThrArgGlyArgSerArgGlyA 439
619 GCGCGCGCGCTGCGCGGATTTTTCGCGCAAACTCGCATACACCAT... 665
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
439 rGserGlyArgSerGlySerPheLeuTyrGlnValSerThrHisGlnGln 455
666 .....GACACTGGCGGCAAAATTCG 685
456 SerGlnSerAlaHisGlyArgAlaGlyProSerThrGlyGlyArgGlnG 472
686 CACACGTCGAAGCGCT 701
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
472 ySerArgHisGlnGln 477
seq_name: sp_human:09H4U3
seq_documentation_block:
ID 09H4U3 PRELIMINARY; PRT; 798 AA.
AC 09H4U3;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, last annotation update)
DE D114N1.1.2 (PROFILAGGRIN 3' END) (FRAGMENT).
GN FLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laird G.;
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL356504; CAC13171.1; -
 DR InterPro; IPR003303; Flaggrln.
 DR PRINTS; PR00487; FILAGGRIN.
 FT MONTER
 SQ SEQUENCE 798 AA: 84773 MW; F923DDA8D1290805 CRC64;

alignment_scores:
 Quality: 142.00 Length: 276
 Ratio: 1.164 Gaps: 10
 Percent Similarity: 44.203 Percent Identity: 21.377

alignment_block:
 us-09-303-518d-571 x 09H4U3 ..

Align seg 1/1 to: 09H4U3 from: 1 to: 798

```

111 GCTGGGAACCGGCTCGACATCTGCGTTTACCTTTAAAGAGAGACC 160
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 SerGlyAspGlySerArgHisSerGly..... 240
161 CGCGCGCATCTCGCCCAATATGCGGCGGCGTTTGAACCCGACAGC 210
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 ....SerHisHisGlnGlnAlaSerSerTrpAlaAspSerArgHis 256
211 CA...CAGCGTCAAGCGCGTTTTCGCGAAAGCGCAAAATCGCGTTTGA 257
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256 eLysSerGlyGlnGlnGlnSerAlaGly..... 265
258 ACTTCCCGCGCGTTTTCAAAAACCGCA..... 287
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 ....ProArgThrSerArgAsnGlnGlnGlySerSerValSerGlnAsp 280
287 ..... 287
280 rAspSerGlnGlnHisSerGlnAspSerGlyArgTrpSerGlySerAla 297
288 ....AGACATCGAAACAGTTTCAAGCGGTACACGCGTGGGAAACAGCT 332
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 eArgAsnHisArgGlySerAlaGlnGlnGlnSerArgAspGlySerArg 313
333 .....GCAGCAGCGTTTGA..... 347
314 HisProThrSerHisHisGlnAspArgAlaGlnHisGlnHisSerAlaG 330
348 CAAGGCGGAAAGCGCTGCTGATCAGCGCGCAACGCGCATCGCAT 396
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
330 uSerSerArgGlnSerGlyThrHisHisAlaGlnAspSerSerGlyGlyG 347
397 TTGGCGGAGCTACATCAGCCAGACTTCGTTCCGTTCCAGTACCGGCA 446
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
347 lAlaAlaSerSerHisGlnGlnAlaArgSerSerAlaGlyGlnArgHis 363
447 GTTCAAGCCGCGCAAAATCAAGCGGTACAGAAATCATGACAGC..... 491
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
364 GlySerHisHisGlnGlnSerAlaAspSerSerArgHisSerGlyLeu 380
492 ....GGCGAGGTCGCGCGCAAGCGCAAAAGCGCGCGCGCATACGA 537
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
380 yHisGlyGlnAlaSerSerAlaValArgAspSerGlyHisArgGlySer 397
538 GGGGT..... 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
397 eArgLysSerGlnAlaSerAspSerGlyGlnHisSerGlnAspSerArg 413
543 .....CAAAACAATCATCA 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

414 GlnSerValSerAlaHisGlyGlnAlaGlyProHisGlnGlnSerHisGln 430
557 AGGCGCTGGCGGCGGCGGAGGACCATCATCTCTCCGACCGACGCTCCCT 606
430 nGlnSerThrArgGlyArg..... 436
607 TCTCCGACGAAGCGCGCGCTGTGGCGGATTTTTCGGCAACCTGC 656
437 ..SerAlaGlyArgSerGlyArgSerGlySerPheLeuTyrGlnValSer 452
657 ATACACCAT.....GACACTGG 673
453 ThrHisGlnGlnSerGlnSerAlaHisGlyArgAlaGlyProSerThrGln 469
674 CGGCAAAATTTGGCACACGTCGCAAGGCGT 701
469 yGlyArgGlnGlySerArgHisGln 478
seq_name: sp_bacteria:Q52747

```

```

seq_documentation_block:
ID Q52747 PRELIMINARY; PRT; 680 AA.
AC Q52747;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ENDO-GUDCANASE.
OS Ruminococcus flavefaciens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Ruminococcus.
OX NCBI_TaxID=1265;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=186;
RA Huang C.M., Asmundson R.V., Yu P.L.:
RT "Nucleotide sequence of a cellulase gene complex from Ruminococcus
RT flavefaciens strain 186 coding for multi-cellulase activities.";
RL Submitted (SEP-1990) to the EMBL/GenBank/DBJ databases.
DR EMBL; X56082; CAA39559.1;
SQ SEQUENCE 680 AA; 75503 MW; 5A589F334846D6F CRC64;

```

```

alignment_scores:
Quality: 141.50 Length: 316
Ratio: 1.004 Gaps: 19
Percent Similarity: 44.620 Percent Identity: 27.848

```

```

alignment_block:
US-09-303-518d-571 x Q52747 ..

```

```

Align seg 1/1 to: Q52747 from: 1 to: 680

```

```

29 CCCCTTTGGAAACGGCATGCATCTGTGACCG...CCCTGCTCAAA 75
   ||| ||| ||| ||| ||| ||| ||| ||| |||
27 ProThnProProserProserThnGlnSerProThnProserSerPr 43
76 TGCTCTCTCC.....TGCTCTGC.....TTTCTG 101
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
43 oAlaSerProArgSerProThnValArgCysArgGlnAlaValAlaPro 60
102 TCTGCACAGCTGGGAAACCGGCTGGGACATCTGGCGTTTACCTTAA 151
   :||| ||||| ||||| ||||| ||||| ||||| |||||
60 lAcys.HisAla.....AspAlaArgThrAlaGly.....Ar 70
152 AGAAGACCGCGCGCATCTGCCAATATGCGCAGCGGTTTGAAC 201
   :||| ||||| ||||| ||||| ||||| ||||| |||||
70 gGlyArg.....Asp 74
202 CCCGACACGACAGCGTCAACCGTTTTCGGAAACGGCAAAATGGG 251
   ||||| ||| :||| ||||| ||||| ||||| |||||
74 roArgHis...GlnCysAspAlaArgLeuArgGlyArgHisArgProArg 89
252 TTTGGAACCTGGCCCGCGCTTTTCAAAAAACGGAAAGCATCGAAACA 301

```

```

seq_name: sp_bacteria:Q9CNR4
seq_documentation_block:
ID Q9CNR4 PRELIMINARY; PRT; 318 AA.
AC Q9CNR4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MSBB.
GN MSBB OR PM0362.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
MDLINE=21145866; PubMed=11248100;

```


RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.
 RT "Complete genomic sequence of Pasteurella multocida pm70."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL: AE006072; AAK02446.1;
 KW Complete proteome.
 SQ SEQUENCE 318 AA; 36768 MW; F91DADB6D9C46375 CRC64;

alignment_scores:
 Quality: 141.00 Length: 268
 Ratio: 0.916 Gaps: 10
 Percent Similarity: 57.463 Percent Identity: 21.269

alignment_block:

US-09-303-518D-571 x Q9CNR4

Align seg 1/1 to: Q9CNR4 from: 1 to: 318

```

112 CTGGGAAACCGCTGGACATCTGGCGTTTACCTTTAAAGGAACCG 161
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 LeuAlaSerTyrIleGlyLysIleValAlaArgLysAlaLysLysGln 69
162 CGCGGCATGTCGCCAATATGCGGAGCGGTTTGAAAC..... 201
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
69 gHisArgAlaLysIleAsnIleGlnTyrCysPheProHisTrpHnGln 86
202 ..CCGACACGACGACGCTCAAGCGCTTTTTCGGAACGCAAAATGC 249
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
86 LAGlnArgLuglnHnTrIleGlnLuglnLuglnPheValIleValHnGln 102
250 GGTTCGAACTGCCCCCGCTTTTCAAAAACCGGAACATGCAAC 299
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
103 MetIleGlyIleGlyLuglnLalaIleArgSerLysArgHnIleGln 119
300 AATGTTCAAGCGGTACACGCTGGACACACGACGAGCGTTTGACA 349
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
119 gArg...SerCysPheHnGlyLuglnIleHisIleHisArgAlaArgLug 135
350 AGCGGAAAGGCTGCTTTCATCACGCGCACATGCGACATGATTTG 399
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
135 LngLysAsnIleIleLeuValProHisAlaTrpAlaIleAspAla 151
400 GCGGAGCGGTACATC..AGCCAGACGCTCCGTTCCACCTGACCGCAT 446
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 SerGlyIleIleLeuHisTrpHnIleGlyMetPro.....MetHis 166
447 GTACAAGCCGCGCAAAATCAAAAGCGATGACAAATCATGACGCGGCA 496
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
166 tYrAsnProHisArgAsnProIleValAspTrpLeuTrpHnIleAla 183
497 GGTGCGGCGCAAGGCAAAACCGCGCCACCGCATACAAAGGGTCAAA 546
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
183 rGlnArgPheGlyLysMetHisAlaArgGln...AsnGlyLys 198
547 CAATCATCAAGGCGCTGCGGCGGCGCAAGCAACATCATCTCCCGCA 596
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
199 ProIleLeuAsnHisValLysGlnIleLysPheGlyTyrTyrLeuPro 215
597 CCACGCTCCCTTCGCGAGAGGCGG.....GGCGTGGGCGG 637
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
215 P.....GlnAspTyrGlyAlaGlnLeuSerValPheValA 227
638 ATTTTTCGCAACCTGCATACACATGACATGCGCGCAAAATTTGCA 687
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
227 sPrPhePheAlaThrTyrLysAlaThrLeuProGlyIleAsnLysIleAla 243
688 CACGTC...AAAGGCGTGAACACCTGTTTCTGCTGCGAAGCGCTGCC 734
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
244 ArgLeuAlaLysAlaAlaValIleProMetPheProArgTyrAsnAlaLe 260
735 CGACGCAAGGCTTCGTTGCACATCCGCGCGCAAGGGAATTTG 784
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
260 uSerGlyLys...TyrGlnIleGlnIleHisProAlaMet...Thrl 275
  
```

seq_name: sp_bacteria:Q51415

```

785 ACGGCACAAAGCCACAGATGCCCGCTGTACACCGCATACCGAATAT 834
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
275 eraSPasPProGlnGlnAlaIleArgAlaMetAsnGlnGlnIleGln 291
835 TGATATACCGCTTTTCGACACAGATATCTGTTATGTACACCGCTATA 884
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
292 PheValThrAlaThrProAlaGlnTyrValTrpIleLeuArgLeuLeu 308
885 AACG 888
308 sThr 309
  
```

seq_documentation_block:
 ID Q51415 PRELIMINARY; PRT; 684 AA.

DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE ALGINATE LYASE.

GN ALG.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=A22 ALG+;

RA Schmitt Andrieu L., Hulén C.;

RT "Cloning and sequencing of a new alginate lyase, algY, from mucoid strains of Pseudomonas aeruginosa, and its expression in E. coli";

RL Submitted (Sep-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z54213; CAA80940.1; -.

KW Lyase.

SQ SEQUENCE 684 AA; 74958 MW; C8CCE638F5675D7A CRC64;

alignment_scores:
 Quality: 140.50 Length: 270
 Ratio: 1.115 Gaps: 10
 Percent Similarity: 46.667 Percent Identity: 26.296

alignment_block:

US-09-303-518D-571 x Q51415

Align seg 1/1 to: Q51415 from: 1 to: 684

```

42 CGCATGCACATCTGTTGAC...CGCGTGCATCAATGCTCTCCCTGC 88
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
304 ArgArgHnHisProPheSerAlaThrProSerAlaThrArgSerProHn 320
89 TGTGCGTTCTCTGTCACACGCTGGGAACCGCTCGACATCTGGCG 138
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
320 rSerSer.....SerGlyLeuAlaIleArgProAlaGlyArgGlyGly 335
139 TTTTACCTTTAAAGGAACGCGCGCGCATGCTGCCAATATGCGGCA 188
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
335 LysLeuProArgGlnAlaHisProArgGlyGlnArgGlnIleAlaLeuSer 351
189 GCGGCGTTTGAACCCGACACGCGACAGCGTCAAAAGCGTTTTCGCGGAA 238
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
352 AlaAlaIleThrAlaValaSerAlaAlaValaLeuArgArgPheLeuGly 367
239 CGGCAAAATGCGGTTTGAACTTGCCCGCGCTTTTCAAAAACCGGAA 288
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
368 .....GlyA 369
289 GACATGCAACATGTTCAACGCGTACACGCGTGGGACAA.....CGT 332
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
369 rGProArgHnGlnArgArgHnGlyGlyArgLeuProAspLeuGlnArg 385
  
```

seq_name: sp_bacteria:Q9F6E3

```
seq_documentation_block:
AC Q9F6E3 PRELIMINARY; PRT; 320 AA.
AC Q9F6E3:
DT 01-MAR-2001 (TrEMBLrel_16, Created)
DT 01-MAR-2001 (TrEMBLrel_16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel_16, Last annotation update)
DE PKSA PUTATIVE O-METHYLTRANSFERASE.
OS Streptomyces collinus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=42684;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM2012;
RT "Putative spore pigment polyketide synthase cluster from Streptomyces
RL submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF293354; AAC26883.1; -.
KW Transferase; Methyltransferase.
SQ SEQUENCE 320 AA; 35453 MW; AEDFA51992JA166F CRC64;
```

```
alignment_scores:
  Quality: 138.50
  Ratio: 1.011
  Percent Similarity: 42.284
  Length: 324
  Gaps: 16
  Percent Identity: 26.235
```

```

alignment_block:
US-09-303-518D-571 x Q9F6E3  ..
Align seg 1/1  to: Q9H6E3  from: 1  to: 320

```

27	TCGCCCTTTGGAAACSSGCAATGCAATCCCTGTT.....GAOCSGCC	67
111111		111
44	SerProValIleProAlaGlyArgLeuIleHisGlyThrProGlyArgAspAlaArg	60
111111		111
68	TCGTAATAGCGCTCCCTCCCTGCTGCTGCTTCCCTGCTGCAACAGCGCTGGGA	117
111111		111
60	GAlaHisArgIleGlyLeuProAlaIleAla.....ArgGlyA	72
118	AAOCSGCTGGCAATCT.....	134
72	rGrProGlnIleProAlaArgIleHisAlaValAlaHisArgAlaValAsp	88
135	...GGCGTTTACCTTTTAAGAAAGAACCGCGCGCATCGTGGCCAAATA	181
111111		111
89	LeuGlyArgLeuAlaGlnAlaGlyArgGlyAlaHisArgProIleArg	105
182	TCGCGCAGCGGGGTTTGAAACCCGCAACGCAAGAGCGTCAAAACCCGTTT	231
105	G.....	105
232	GGGGAAGACGGCAAAATGGCGGTTTGGAACTGGCCCGCGCTTTTCAAAA	281
111111		111
106	..GlyGlyIleProAlaIleArgIleIleValLeuProLeu.....	117
282	ACCGGAACACATCGAAACATGTTCAAGCGGCTACACGCGTGGGAACAGC	331
118ProGlnArgGlyArgProGlyValGlyValArgLeuIleProArg	132
332	T.....GCAGCAGCGCTTTGGCAACAGGCGCAAGCGCT.....	362
111111		111
132	GHisAspArgLeuGlnArgAlaValGlyAlaGlyAlaArgArgValProAla	149
363GCTGTATATCAACCGCGCATATGCGCAGCTACGATTGTGGCGG	404
149	rGrProHisGlyMetHisLeuIleGlyIleArgArgAlaArgArgProGlyAla	165
405	ACGTAATCATCAGCAGCAGCTTCGCTGCA...CTGACCGCCATGTACA	451
166	ArgArgGlyGlnProAlaArgGlyValArgProIleProAlaArgAsp**Al	182
452	AGCGCGGCAAAATCAAAACGATAGCAAAATCATGACGCGGCGAGCGAT	501
182	aArgProArgProGlyGlyGlyIleGlyIleGlyArgArgProGlyThrAlaHisGlyA	199
502	CGCGGCAAAAGCAAAACCGCGCCGACGCGCATACAGGGGTCAAAACAAT	551
111111		111
199	rGrArgProArgIleProArgAlaArgArgAlaArgArgProAspGly	215
552	CATCAAGGC.....CCTGGCGCGG	571
216	ArgProGlyGlyGlyArgArgLeuArgAspGlnIleHisProGlyValGly	222
572	GGCAGGACCAATTCATCTCT.....GCCGACCAACGTCCTCTTCG	612
111111		111
232	YArgArgGlyIleHisIleHisProHisAlaIleGlnArgProArg.....	245
613	CAGGAAGCGCGCGGCT.....GTGGCGGATTTTTCGGCAAACTGC	656
111111		111
246	..GlyArgArgProArgArgAspAlaArgGlyGly.....	255
657	ATTCACACATGACACAGCGCGGCAAAATTCGCG.....ACAGTCAAGCG	700
111111		111
256	...HisArgGlu**ArgArgArgLeuAlaLeuAspAlaValGlnHisArg	271
701	TGAAAACCTGTGTTTTGCTGCGAACGCGTCCGCGACGCAAGCGTTC	750
111111		111
271	gAspGlyPro.....AlaAlaIleArgProGlnArgArgArg	283

751 GTGTGCACATGCGGGCCGGTCCAAAGSGCAATTGAACGGCACAACAGGCCA 800
||||||| ||| |||
283 TgcInalIahIsProAlaLglnHisglYcylPProAlaIahIsArglglyAtgPro 299
||||||| :|||
801 CGATGCGCGCGCGTTCCAACGGCAATACGCAATATTGGATTCGCGCGTTTC 850
||||| ||||| ||||| M |||||
300 GlynHisArgTargHISglInPro.....,glYcInProLeuSe 311
|||||||
851 CGACGCGCATATCTGTTTAATA 872
| :||| :|||
311 rAlaArgrValarLgLeuHisIle 318

seq_name: sp_bacteriap:Q9A6A1

```

seq_documentation_block:
ID AC 09A6A1 PRELIMINARY; PRT; 424 AA.
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN CC2193.
GN CC2193.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_taxonomy:69394;
XN [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed-11259647;
RA Nieman W.C., Feldhlyum T.V., Taub M.T., Paulsen I.T., Nelson K.E.,
RA Elsen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Pollock A.I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Craven M.B., Kouri H., Haft D.H.,
RA Kolonay J.F., Smit J., Gwin A.S., Gwin M.L., Shetty J., Barry K.,
RA Salzberg T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Frazer C.M.,
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AEO05891; AAK24164.1; -.
DR TIGR; CC2193; -.
DR InterPro; IPR002048; EF-hand.
DR PROSITE; PS00018; EF_HAND; UNKNOWN.2.
KW Hypothetical protein; Complete proteome.
SQ
SEQUENCE 424 AA; 47293 MW; 4E1FED040A22F933 CRC64;

```

```
alignment_scores:
```

Quality:	138.00	Length:	305
Ratio:	0.986	Gaps:	18
Percent Similarity:	45.902	Percent Identity:	27.213

alignment_block:

Align seg 1/1 to: Q9A6A1 from: 1 to: 424

```

30  CCGTTCCTGGAAACCGCATCATGCTCGTGTACGCCCTGTCGAATGCC  79
   |||||:::|||||
5  ProhevalSerArg.....ArgProspaArgArg  15
   |||||
80  TCTCCCTGCTGTCGCTTCTCTGTCACACGCTGGAAACGGCT....  126
   |||||
15  galArgProAla.....ArgArgAlaArgAlaArgProArgPro  29
   |||||
126 .....CGGACATGCGCGTTTACCTTTAAAGCA  155
29  heaAlaLeuSerValAlaAspArgValArgCysValArgPro...GIgArg  44
   |||||
156  AGACCGCGGCGCG.....CATGCTGC...CAATATGCGGACGGCG  193
   |||      |||||:::||||
45  ArgHisGlyAlaGlyGlyLeuLeuHisArgGluLeuLeuHisAlaAlaGlyGly  61

```

[illegible]


```

265 .....CCCCGTTTTCAAAAACC 284
105 LeuIleThrIleValThrSerSerArgAsnProLysGlyPheSerSergl 121
285 G.....GAAGACATGCAACAATGTTCAAGCGGTACACG 319
121 uGUValIleSerAsnGluAspLeuGlnIleThrPheLys..... 134
320 GCTGGACACGTGCAGCGCTTGGCAAGCGGAGCGGTCTGCTTC 369
135 .....AsnLeuGlnIleLysGlnIleLeuLeu 144
370 ATCAGCGCCGACATCGCGACGTACGATTGGCGGACGCTACATCAGCA 419
145 PheCysGlyHisGlnAlaAsnTrpGluLeuProPheLeuTyrIleThrLys 161
420 GCAGCTCCGCTCCACCTGCAGCGCATGTACACGCCCGCAAAATCAAG 469
161 sAsnTyrProGly.....IleAlaPheAlaLysAlaIleLysAsnGlnA 176
470 CGATAGACAAATCATGCGCGGCGGCGGCGGCAAGGCAAAACC 519
176 rGleuSerLysLysIlePheAlaLeuArgGluValPheLysGlyLysIle 192
520 GCGCCCGACCGCATACAAAGGGGTCAACAATCATCAAGCGCTCGCGC 569
193 ValProProlLys...AsnGlyIleGlnGlnGlyIleGlnAlaLeuAsnGln 208
570 GGGCGAGCAACCATCATCTGCGCGACGCTC.....603
208 nGlyLysLeuValGlyIleValGlyAspGlnAlaLeuLeuMetSerSergl 225
604 .....CCTTCTCCGACGAGGAGCGCGCGCTGTGGCGATTGTTTC 645
225 yTrhTyrPro.....LeuPhe 230
646 GCGAAACCTGCATACCATGACATGCGCGCAAAATGCGACACGTCAA 695
231 GlySerProAlaPheThrThrThrSerProAlaLeuAlaTyrLysTh 247
696 AGCGGTGAACCCCTGTTTCTGCTGCGAACCCCTCGCGACGACAG 745
247 rGlyPheProValIleAlaValAsnValSerArg.....GlnAlaLysG 262
746 GCTTGTGTTCACATCGCGCCCGTCCAGAGGGAATTGAACGCAACAA 795
262 Lyphe.....GluValIlePro...SerAlaLysLeuTyrAlaAsnLys 275
796 GCC.....CAGGATCGCGCGCTGTTCACCGCAAT.....825
276 SerLeuProMetLysGlnSerValAlaIleLeuMetAspGlnMetMetGln 292
826 .....ACCGAATATTGGATACGCCGTTTCCGACGACGATCTGTATTGT 871
292 yPheLeuGlnLysGlyIleAlaSerGlnProGlnIleTrpMetIleH 309
872 ACAACCGCTATAA 885
309 IsLysArgTrpLys 313

```

seq_name: sp_bacteriap:092983

```

seq_documentation_block:
ID 092983      PRELIMINARY;      PRT;      467 AA.
AC 092983;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE ACYLTRANSFERENCE.
GN HRRB OR CPN0098.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
NCBI_TaxID=83558;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CW1029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
DR EMBL; AF001596; AAD18251.1; -.
KW Transferase; Complete proteome.
SO SEQUENCE 467 AA; 53193 MW; D3C7C284E9220DD0 CRC64;

alignment_scores:
Quality: 137.00      Length: 338
Ratio: 0.867      Gaps: 15
Percent Similarity: 46.746      Percent Identity: 23.964

alignment_block:
US-09-303-518D-571 x 092983

Align seg 1/1 to: 092983 from: 1 to: 467

52 ATCTGTTGACCGCCCTCTCAAAATGCTCTCCCTGCTGCTTCCTG 101
||||| :||| :||| :||| :||| :|||
12 IleLeuGlnAlaProLeuTyrTyrLeuValSerGlyIleIleLeuLys 28
102 TCTGCACAGC.....CTGGCAACCGCGCTCGGAC 130
| :||| :||| :||| :||| :||| :|||
28 sArgHisThrProArgSerPheLeuThrGlyLeuGlyLysGlyPheGly 45
131 ATCTGCGTTTACCTTTTAAAGAGCGCGCGCGCATCTGCGCAAT 180
||||| :||| :||| :||| :||| :|||
45 IleLeuAlaPheTyrIleIleSerAspTyrArgLysThrAlaLeuThrAsn 61
181 ATGCGGAGCGGGGTTTGAACCCCGACGACGAGCGTCAACCCGTTT 230
||| :||| :||| :||| :||| :|||
62 LeuAlaLeuAla.....PheProGlnLysThrHrp 71
231 TCGCGAAGCGCAAAATGCGCT.....252
||| :||| :||| :||| :||| :|||
71 eAspGluArgTyrLysIleAlaArgIleSerLeuGlnHisLeuIleIle 88
253 .....TTGAACCTGCC.....264
||||| :||| :||| :||| :||| :|||
88 hIleLeuGlnIleLeuAlaIleGlnLeuValLysnIleAspLys 104
265 .....CCGCGCTTTCAAAAACC 284
105 LeuIleThrIleValThrSerSerArgAsnProLysGlyPheSerSergl 121
285 G.....GAAGACATGCAACAATGTTCAAGCGGTACACG 319
121 uGUValIleSerAsnGluAspLeuGlnIleThrPheLys..... 134
320 GCTGGACACGTGCAGCGCTTGGACAGGCGGCAAGCGCTGCTGTC 369
135 .....AsnLeuGlnIleLysGlnIleLeuLeu 144
370 ATCAGCGCCGACATCGCGACGTACGATTGGCGGACGCTACATCAGCA 419
145 PheCysGlyHisGlnAlaAsnTrpGluLeuProPheLeuTyrIleThrLys 161
420 GCAGCTCCGCTCCACCTGCAGCGCATGTACACGCCCGCAAAATCAAG 469
161 sAsnTyrProGly.....IleAlaPheAlaLysAlaIleLysAsnGlnA 176
470 CGATAGACAAATCATGCGCGGCGGCGGCGGCAAGGCAAAACC 519
176 rGleuSerLysLysIlePheAlaLeuArgGluValPheLysGlyLysIle 192
520 GCGCCCGACCGCATACAAAGGGGTCAACAATCATCAAGCGCTCGCGC 569
||||| :||| :||| :||| :||| :|||

```

193 ValProProlys...AsnGlyIleGlnGlnGlyIleGlnIleValLeuAsnG1 208
 570 GGGGAGGCAACCATCATCCTGCCGACACGTC..... 603
 208 nGlyLysLeuValGlyIleValGlyAspGlnIleLeuLeuMetSerSert 225
 604CCTTCTCCGAGAGAGGCGGCGCTGTGGCGGATTTTTC 645
 225 YrThrTyPro.....LeuPhe 230
 646 GGCAACCTGCATACCATGACACTGGCGCAAAATTGGCACACGTC 695
 231 GlySerProAlaPheThrThrSerProAlaLeuAlaTyrLysTh 247
 696 AGGCGTGAACCCCTGTTTCTGCTGGACAGCCCTGCCGACGACAG 745
 247 rGlyPheProValIleAlaValAsnValSerArg.....GlnAlaLysG 262
 746 GCTTCGTGTGCACATCCGCCCTGCCAAGGGGAATTGAACGGCAACA 795
 262 LyrPhe.....GluValIlePro...SerAlaLysLeuTyrAlaAsnLys 275
 796 GCC.....CAGCATGCCGCTGTTCACACCGCAAT..... 825
 276 SerLeuProMetLysGlnSerValAlaIleLeuMetAspGlnMetMetG1 292
 826 ...ACGGAATTTGGATACGCCGCTTTCGACGACGATCTGTTATG 871
 292 yPheLeuGlnLysGlyIleAlaSerGlnProGlnGlnIleThrPhe 309
 872 ACACCGCTATATA 885
 309 IsLysArgTrpLys 313
 seq_name: sp_verticbrate:Q91238

seq_documentation_block:
 ID Q91238 PRELIMINARY; PRT; 326 AA.
 AC Q91238;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE HIGH MOLECULAR WEIGHT BASIC NUCLEAR PROTEIN (FRAGMENT).
 OS Pseudopleuronectes americanus (Winter flounder) (Pleuronectes americanus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Acanthopterygii; Perciformes; Pleuronectiformes; OC Pleuronectidae; Pleuronectidae; Pseudopleuronectes.
 OX NCB1 TaxID=8265;
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RX MEDLINE=98165790; PubMed=9497335;
 RA Watson C.E., Davies P.L.,
 RT "The high molecular weight chromatin proteins of winter flounder sperm are related to an extreme histone H1 variant."
 RL J. Biol. Chem. 273:6157-6162(1998).
 RN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 RX MEDLINE=99196707;
 RA Watson C.E., Davies P.L.,
 RT "Recent and rapid amplification of the sperm basic nuclear protein RT genes in winter flounder."
 RL Blochim. Biophys. Acta 1444:337-345(1999).
 DR EMBL: U39735; AAC13878.1;
 KW Nuclear protein.
 FT NOWITER 1
 SQ SEQUENCE 326 AA; 38103 MW; 83E3D5B773EC3F77 CRC64;

alignment_scores:

Quality: 133.00 Length: 261
 Ratio: 1.064 Gaps: 11
 Percent Similarity: 47.893 Percent Identity: 26.820
 alignment_block:
 US-09-303-518d-571 x Q91238

Align seg 1/1 to: Q91238 from: 1 to: 326

164 CGGCATCGTCGCCAATATGCGGACGCGGTTTGAACCCGACACGCG 213
 83 ArgSerLysSerPro.....LysArgArgValGlnThrProlLysMetArg 97
 214 ACGGTCAAAACCGTTTTCGCGAAGACGCAAAATGCGTTTGAACCTGC 263
 97 galThrSerPro...MetArgSerArgSerProSerArgTyrLysSerP 113
 264 CCCCCGTTTTCAAAAACCGGAGACATTCGAATGTCACAAAGCG 313
 113 rGlySerValLysThrProlLysThrArgAlaAsnSerAlaArg 129
 314 TACACGCTGGAAACAGTCACAGCTTTGGACAAAGGCGGCTG 363
 130 ProlLys..... 131
 364 CTGTTATACGCGGCACATCGGCACCTACATTTGGCGGACGCTACT 413
 132 ...SerProMetArgSerProSerProlLysMetArgAlaLysThrArg 147
 414 CAGCCAGACCTTCGTTCCACTGACGCGCATGTCACAGC..... 454
 147 LysArgSerProlLysArgSerAsn...ProProlLysArgSerValLysThr 162
 455CGCCGAAATCAAGCGATAGACAAATATCATGACGCGG... 493
 163 ProlLysThrArgAlaLysSerProArgArgSerLysSerProlLysArg 179
 494GCAGGCTGCGGCGCAAGGCA 515
 179 gValGlnThrProlLysMetArgAlaLysSerProMetArgSerArgLys 196
 516 AACCGGCGCCACCGCATACAGGCGTCAACAAATATCATGACGCGCTGC 565
 196 rGlnArgSerProSerArgSerThrSerProMetArgSerGlnSerPro 212
 566 GCGGCGGCGAGGACCATCATCTGCGGACGACCTCTCTCCGAG 615
 213 LysArgThrValLysArgGlnLysMetThrAlaLysSerLeuMetArg 229
 616 GAAGCGGCGGCGTGGCGGATTTTTCGCAAACTGCATACACCAT 665
 229 rArgSerProSerArgSerLysSerProlLysArgSerValLysThrPro 245
 666 GACACTGCGGCAAAATTTGGACACGCT.....CAAAGCG 700
 246 LysThrArgAlaLysSerProArgArgSerLysSerProlLysArgVal 262
 701 TGAACACCTGTTTTCGCTGCGAAGCGCTGCCGACGAGCA..... 743
 262 LglnThrProlLysArgArgValGlnThrProlLysArgValGlnThrP 279
 744AGGCTGCTGTGCACATTCGCGCCCG 770
 279 rGlySerArgLysProlLysProlLysArgArgValProThrProlLysArg 295
 771CAAAGGGGAATTTGAAGGCAACAAGCCACGATGCGCGTGT 814
 296 ArgValProlLysProlLysArgArgValGlnThrProlLys...ArgArgVal 311
 815 TCAACCGCAATACGATATTTGATACGCG 845
 311 LglnThrProlLysArgArgLeuGlnThrPro 321

seq_name: sp_bacteriap:09RYP2

seq_documentation_block:

ID 09RYP2 PRELIMINARY; PRT: 376 AA.

AC 09RYP2:

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE ADENINE DEAMINASE-RELATED PROTEIN.

GN DRA0268.

OS Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.

OX NCBI_TaxID=1299;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RI;

RX MEDLINE=2003696; PubMed=10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson K.J., Hatt D.H., Gwin M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,

RA Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus

RL radiodurans R1."

RL Science 286:1571-1577(1999).

DR EMBL; AE001863; AAF12376.1;

DR TIGR; DRA0268;

KW Complete proteome.

SQ SEQUENCE

376 AA; 39845 MW; 7AB7FF32P8C45651 CRC64;

alignment_scores:

Quality: 132.00

Length: 351

Ratio: 0.852

Gaps: 21

Percent Similarity: 44.160

Percent Identity: 26.456

alignment_block:

US-09-303-518d-571 x 09RYP2

Align seg 1/1 to: 09RYP2 from: 1 to: 376

```

18 CAGCGTGTTCCTCCCTTTCGCAACCGCATGACATCTGTGACCGCC 67
   ::::::::::: :::::::::::
23 ArgAlaAlaAspGluLeuGlyAsnArgCysArgHis.....Argpr 36
68 TCGTAAATGCTCTCCCTGCTGCTGCTCTGCTGCTGCTGCTGCTGCT 113
   ||| ::| ||| :::::::::::
36 Ala...LeuArgGlyAspAspSerAlaArgLeuAlaGlyHisAlaHisH 52
114 .....GGAAACCGGCTCGACATCTGCGCTTTTACCTTTTAA 152
   ||| ::| ||| :::::::::::
52 IstHrHisPheGlyAlaAlaGlyArgThrGly.....Thr 63
153 GGAAGACCGCGCGCGCATCGTCCCAATATGCGGACGCGGTTTGAACC 202
   ||| ||| ||| ||| |||
64 GlyArgProAlaAlaGlyProAlaGlyGlnSerGlyAlaGlyGlnPr 80
203 CGGA.....CAGCAGACGCTCAAGCGCTTTT 231
   ||| ::| ||| :::::::::::
80 OArgArgValAlaHisProAlaLeuTyrSerGlyGlyArgAlaAspArg 97
232 GCGGAACGCGCAAAATGCGGTTTGGACATTCGCCCGCGGTTTTCAAAA 281
   ::| ||| ||| ||| |||
97 IagIyArgAlaGlySerAlaAlaGlyAlaArgTyrArgGluIleLeuGly 113
282 ACCGGAAGACATCGAAACAAT.....GTTCAAG 310
114 ProAlaArgProAlaGlyAspArgProThrValSerArgArgLeuArgAl 130
311 CGGTACACGCGCTGGGACA...CGTGCAGACGCTTTGGAACAAGGCGAA 357
   ::| ||| ||| ||| ||| |||

```

```

130 AlaAlaAlaArgLeuAlaThrLeuArgGlyAlaGlyAsnLeuArgGlyArg 147
358 GGGCTGCTGTTCATCAGCGCCGACATCGCGAG..... 389
   ||| ::| ||| ||| |||
147 rglYser.....AlaAlaArgArgArgAspAlaProProAlaArg 160
390 .....CTAGATTTCG 400
161 ThrArgArgArgProAlaAlaArgLeuGlyArgGlyHisLeuArgSerAl 177
401 GCGGACGCTA...CATCAGCCAGCGCTCCGTTCCATCGACCGCCATG 447
   ||| ::| ||| ||| |||
177 ArgAlaLeuAlaHisAlaProAlaAspValSer...ArgProAspArgHis 193
448 TCAACGCGCGCGCAAAATCAAGCGATGACAAATCAATCAAGCGC..... 491
   ::| ||| ||| |||
193 rGAlaCysGlyAlaGlyGln.....ArgArgArgProAlaGlyArgArg 207
492 .....GGCGAGG 499
208 ArgProLeuArgAlaAlaArgValAlaGlyLeuLeuAspValGlyGlnAr 224
500 TCGCGCGCAAGG.....CAAAACGCGCGCCACCGGCATACAA 537
   ::| ||| ||| |||
224 gProAlaArgArgHisProGlyHisGlnHisSerAlaArgArgAlaSerG 241
538 GGGGTCAACCAATCATCA.....GCCCTGCGCGCGG..... 572
   ||| ::| ||| ||| |||
241 IyGlyProAlaAlaArgArgGlnArgArgGlyProAlaArgGlyGlySerGly 257
573 .....CGAGCAACCATCATCTGCGCGA.....CC 598
   ||| ::| ||| ||| |||
258 AlaArgThrAlaArgArgArgHisArgProGlyArgGlyArgGlyArg 274
599 ACGTCCCTTC.....TCGCGAG 615
   ||| ||| ||| |||
274 OArgAlaValAlaAlaProAlaLeuArgGlySerAspAspArgProAlaSerG 291
616 GAAGCGCGCGCGCTGCGGATTTTTCGCAACCTGCATACACAT 665
   ||| ||| ||| |||
291 IyArgGlyArgArgCysProGly.....ProGly 300
666 GACACGTGCGCGCAAAATTCG.....ACAGTCAAGCGGTGAAA 706
   ||| ||| ||| |||
301 AspGlyGlyAlaLeuAlaGlyLeuHisLeuAlaLeuSerArgHisHis 317
707 CCGCTGTTTCTG.....CTGC 723
   ||| |||
317 sProGlyHisSerArgProGlyArgAspSerGlyAlaGlnAlaAspAla 334
724 GAAGCGCTGCGCGCGCAAGGCTTGTTCACATTCGCCCGCCGCCA 773
   ||| ||| ||| |||
334 IaaArgProAlaArgArgAspGlyValAlaAlaLeuAlaGlyProGlyArgVal 350
774 AGG 776
   |||
351 Arg 351

```

seq_name: sp_bacteria:Q53478

seq_documentation_block:

ID Q53478 PRELIMINARY; PRT: 456 AA.

AC Q53478:

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE ORF2 PROTEIN.

GN ORF2.

OS Neisseria gonorrhoeae.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

OX NCBI_TaxID=485;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-95020543: PubMed-7934834;
 RA Facius D., Meyer T.F.:
 RT "A novel determinant (coma) essential for natural transformation
 RT competence in *Neisseria gonorrhoeae* and the effect of a coma defect on
 RT pillin variation.";
 RL Mol. Microbiol. 10:699-712(1993).
 DR EMBL: S75490; AAB32262.1; -;
 SQ SEQUENCE 456 AA; 54380 MW; 30AC7B5CDFB912D5 CRC64;

alignment_scores:
 Quality: 130.50 Length: 287
 Ratio: 0.981 Gaps: 17
 Percent Similarity: 46.341 Percent Identity: 26.829

alignment_block:
 US-09-303-518D-571 x Q53478 ..

Align seg 1/1 to: Q53478 from: 1 to: 456

```

126 CGGACATCGCGCTTTTACCTTT.....AAGAGAACCGCCG..... 164
    ||||| ||| ||| ||||| |||||
170 ArgThrValGlyTyrLysProGlnGlnHisArgProArgThrG1 186
    ||| ||||| ||| ||||| |||||
165 .....GGCATCGCGCCATATGCGGAGCGCGGTTTGACCCGAC 207
    ||| ||||| ||| ||||| |||||
186 nhrGlnAlaAlaArgGlnGlnGlnGlnSerGlyThrHisArgGln 203
    || ||| ||| ||||| ||||| |||||
208 AGCAGACGCGTAAAGCGCTTTTGGGAAACGCGCAAAATGCGG..... 251
    || ||| ||| ||||| ||||| |||||
203 isGlnGlnArgGln...ArgCysGlyGlyAsnGlyLysLeuArgGlnHis 218
    || ||| ||| ||||| ||||| |||||
252 TTTTGA.....ACTTGCC 265
    |||||
219 IIEGlyLysProHisGlnLysProGlnSerValPheArgGlnLysCysP 235
    |||||
266 CGCGCTTTTCAAAAACCGAAGACATCGAACAATGTTCAACCGGTA 315
    ||||| ||||| ||||| |||||
235 oArgLeuLeuLysArgArgGlyArgHisArgThrGlnGlnArgGlnA 252
    |||||
316 CA.....CGCGTGGGAAACGACT 332
    |||||
252 rGgLyGlnHisProArgLysProArgTyrArgHisArgLLeuSerGln 268
    |||||
333 GCAGCAGGCTTTGGACAGGCGGAGGCTGTTCAAC..... 374
    || ||| ||||| ||||| |||||
269 AlaAspGlnArgGlnArgGlyLysGlnLLeuGlnGlnGlnArgGln 285
    |||||
375 .....GCCGACATCGCGACGCTACGATTTGGCGGA 405
    |||||
285 nHisArgSerArgProLeuProAlaHisGlyLysLeuProPhe..... 299
    |||||
406 CGGTACATGACGACGCTCGCTCCGCTCCGACCGCGCATACAGGC 455
    |||||
300 .....ProLeuValGlnThrProCysThrArgProTyrGln... 311
    |||||
456 GCCGAATATCAAGCGATAGACAAATATCATGCAAGCGGAGGATGCGG 505
    ||||| ||||| ||||| |||||
312 GlyArgHisGlnThrGlnArgGlnProSerThrAspThrGlnAsn... 327
    |||||
506 GCAGAGCGCAAAACCGCGCCACCGCATCAAGGGGTCAACAAATCATC 555
    ||||| ||||| ||||| |||||
327 gGlnArgValGlnGlnGlnHisArgArgGlnArg.....L 339
    |||||
556 AAGGCGCTCGCGCGGCGGAGGAGACATCATCGCCGACGACGATGCC 605
    ||| ||| ||| ||| ||||| |||||
339 euProProArgArgProArgArgGlnSerSerProAla...ProArgPro 354
    |||||
606 TTCTCCGAGGAAGCGCGCGCTGTGGCGGATTTTTCGCAAAACGTC 655
    ||||| ||||| |||||
355 .....ArgGlnArgArgGln..... 359
    |||||
656 CATACACCATGACATGCG.....GCCAAATTTGGCACACGTC 693
    |||||
  
```

```

360 .....HisGlnAsnThrAlaLeuArgArgHisGlyLysThrGlyLysGlnA 375
    ||||| ||||| |||||
694 AAGGCGGTGAAACCGCTGTTTCTGCTCGCAACGCGCTCCGACGAGACA 743
    ||||| ||||| ||||| |||||
375 rGValLysGlnArg.....AlaArgThrPro...ArgArgGln 386
    |||||
744 AGGCTTCGTTCGATCGACATCGCCCGCTCCAGAGGGAATTTGAACGGCA 793
    ||||| ||||| ||||| |||||
387 AsnProArgGlnArgProGlyGlyLysAlaArgGlyThrGlyLysGlnPr 403
    |||||
794 AAGCCACAGATGCGCGCGTGTTCACCCGCAATACGCAATTTGATAGCC 843
    ||||| ||||| ||||| |||||
403 oLeuArgGlnProArgGlyGlnHisArgHisHisArgAsnValGlnThrA 420
    |||||
844 CGTTTCCGAC 854
    |||||
420 rGAsnThrAsp 423
    |||||
  
```

seq_name: sp_bacteria:Q93LB1

```

seq_documentation_block:
ID Q93LB1 PRELIMINARY; PRT; 299 AA.
AC Q93LB1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE DNA-BINDING PROTEIN.
OS Bifidobacterium animalis (Bifidobacterium lactis).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Bifidobacteriales; Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=28025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27536;
RA Gonzalez Vera A., Rossi M., Altomare L., Gotti R., Maltouzzi D.;
RT "Effect of pH and culture redox potential on the continuous culture of
RT Bifidobacterium animalis ATCC 27536, studies on stability of
RT recombinant plasmids and characterization of the insertion sequence
RT IS199.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ318089; CAC41624.1; -.
KW DNA-binding.
SQ SEQUENCE 299 AA; 34340 MW; B71D96C9EC0A7A03 CRC64;
  
```

alignment_scores:
 Quality: 130.00 Length: 281
 Ratio: 1.111 Gaps: 15
 Percent Similarity: 41.637 Percent Identity: 26.335

alignment_block:
 US-09-303-518D-571 x Q93LB1 ..

Align seg 1/1 to: Q93LB1 from: 1 to: 299

```

105 GCACACGCTGGGAAACCGCTCGACATCTGCGGTTTACCTTTAAAG 154
    ||||| ||||| ||||| |||||
16 AlaArgAlaGlyArgProVal...LeuLeuGlyAlaArgAspArgG1 31
    |||||
155 AAGACCGCGC..... 164
    |||||
31 YArgAlaArgValAlaProArgArgValArgArgProValArgAspArgA 48
    |||||
165 .....GGCATCGTCGCCAA..... 179
    |||||
48 sPArgArgProAspProAlaLeuArgArgGlnAlaGlyLysThrAlaGly 64
    |||||
180 TATCGCGAGCGGGGTTTGAACCCGACACGACGACGCGTCAAGCGTTT 229
    ||||| ||||| |||||
65 HisGlyAlaGlyProPheArgPro..... 72
    |||||
230 TTGGCGAAACGCAAAATGCGGTTTGAACCTTGCCCGCGGTTT... 275
    |||||
  
```

RL Mol. Microbiol. 29:629-639(1998).
 RN [2]
 RP SEQUENCE OF 534-1122 FROM N.A.
 RC STRAIN-MC215;
 RC MEDLINE=95191405; PubMed=7885234;
 RA Fiss E.H., Yu S., Jacobs W.R. Jr.;
 RA "Identification of genes involved in the sequestration of iron in
 RT mycobacteria: the ferric exochelin biosynthetic and uptake pathways.";
 RL Mol. Microbiol. 14:557-569(1994).
 RN [3]
 RP SEQUENCE OF 534-1122 FROM N.A.
 RC STRAIN-MC215;
 RC MEDLINE=98389687; PubMed=9721311;
 RA Yu S., Fiss E., Jacobs W.R. Jr.;
 RT "Analysis of the exochelin locus in *Mycobacterium smegmatis*:
 RT biosynthesis genes have homology with genes of the peptidase
 RT family.";
 RL J. Bacteriol. 180:4676-4685(1998).
 DR EMBL: AF034152; AAC32046.1; -;
 DR EMBL: AF027770; AAC82548.1; -;
 DR InterPro: IPR001140; ABC_transport_tmam.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00664; ABC_membrane.1.
 DR Pfam: PF00005; ABC_tran.1.
 KM Hypothetical protein.
 SQ SEQUENCE 1122 AA; 120386 MW; 4EC826C67914A57A CRC64;

alignment_block

3 GYTCGTTACAAATTCAGGCTGTTCCCCCTTGGGAACCGGCATGCACA 52

[illegible]

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ohtaki S., Umeki K., Sawada Y.;
 RT "Homo sapiens mRNA for RNA binding protein, partial cds."
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB016088; BA83714.1;
 DR InterPro; IPR002965; P-rich_extensn.
 DR PRINTS; PR01217; PRICHEXTENSN.
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 956 AA: 106462 MW: 9764078495011 CRC64;

alignment_scores:

Quality: 128.00 Length: 260
 Ratio: 1.153 Gaps: 11
 Percent Similarity: 42.692 Percent Identity: 26.154

alignment_block:

US-09-303-518D-571 x Q9UQ39 ..

Align seg 1/1 to: Q9UQ39 from: 1 to: 956

```

152 AGAAGACCGCGCGCATGTCGCCAATATGCGGAGG..... 190
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
444 ArgArgMetGlyArgSerArgSerProGlnArgGlyArgSer 460
191 .CGGGTTGAACCGCGACGACGAG.....CGGTCAAGC 224
|||||:|||||:|||||:|||||:|||||:|||||:|||||
460 ArgSerArgThrProGlnArgGlyHisSerArgSerArgSer 477
225 CGTTTTCGGAACCGCAAAATCGGTTTGAACTTGCC..... 265
|||:|||||:|||||:|||||:|||||:|||||:|||||
477 IntParArgSerArgSerArgSerArgGlnArgTrpGlyArg 493
266 .....CCGCGTTTTCAAAAA 282
|||||:|||||:|||||:|||||:|||||:|||||:|||||
494 GlnArgArgGlyArgSerArgSerProGlnArgProGlyTrp 510
283 CCGGAAGCATCGAACAATGTCAAGCG..... 313
|||||:|||||:|||||:|||||:|||||:|||||:|||||
510 ArgGlnThrGlnArgArgGlyArgSerArgSerAlaArgGly 527
314 .....TACAGGCTGGGAACAGTCGACGAGCGCTTTGG 346
|||||:|||||:|||||:|||||:|||||:|||||:|||||
527 erHisSerArgSerProAlaThrArgGlyArgSerArgSer 366
347 ACAAGGCGAAGCGCTGCTTCATCAAGCGCACATGCGCAGT 396
|||||:|||||:|||||:|||||:|||||:|||||:|||||
542 ThrProAlaArgGlyArgSerArgSerArgThrProAla 555
397 TTGGGCGGACGCTACATCAGCCAGCAGCTTCCGCTGACGCG 446
555 ..... 555
447 GTACAAGCGCGCAAAATCAAGCGATAGCAAAATCATGCGG 496
|||||:|||||:|||||:|||||:|||||:|||||:|||||
556 .....ArgArgArgSerArgSerArgThrProAlaArg 568
497 GGGTGGCGGCAAGGCAAAACCGCGCCACCGCATACAAAGG 546
|||||:|||||:|||||:|||||:|||||:|||||:|||||
569 .....SerArgSerArgThrProAlaArgGlyArgSe 580
547 CAATCATCAAGCGCTGCGCGCGGAGGCG.....AACCAT 593
|||||:|||||:|||||:|||||:|||||:|||||:|||||
580 ArgSerArgThrProAlaArgArgArgSerArgThrProAl 596
594 CGACCGACGCTCTTCGCGAGAGGCGGCGGTGGCGGATTTT 643
|||||:|||||:|||||:|||||:|||||:|||||:|||||
597 ArgArgArgSerArgSerArgSerProAlaArgSerGly 610
644 TCGGCAAACTGTCATACACATGACATGCGCGCAAAATTGG 693

```

```

611 .....Arg 612
694 AAGCGGTGAACCCCTGTT.....TTTGTGCGGACGCGCT 731
|||||:|||||:|||||:|||||:|||||:|||||:|||||
612 erArgSerArgThrProAlaArgArgGlyArgSerArgSer 628
732 GCCCGACGACAGCGCTGTTGCACATCGCGCCCGCAAGGG 781
|||||:|||||:|||||:|||||:|||||:|||||:|||||
629 AlaArgArgGlyArgSerArgSerArgThrProAlaArg 645
782 TGAACGGCAACAAGCCGACGATGCCG 809
645 gSerArgSerArgThrProAlaArgArg 654

```

seq_name: sp_human:Q9UQ40

seq_documentation_block:

```

ID Q9UQ40 PRELIMINARY; PRT; 1262 AA.
AC Q9UQ40;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RNA BINDING PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohtaki S., Umeki K., Sawada Y.;
RT "Homo sapiens mRNA for RNA binding protein, partial cds.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016088; BA83713.1;
DR InterPro; IPR002965; P-rich_extensn.
DR PRINTS; PR01217; PRICHEXTENSN.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 1262 AA: 138003 MW: 75676BD73FA14295 CRC64;

```

alignment_scores:

Quality: 128.00 Length: 260
 Ratio: 1.153 Gaps: 11
 Percent Similarity: 42.692 Percent Identity: 26.154

alignment_block:

US-09-303-518D-571 x Q9UQ40 ..

Align seg 1/1 to: Q9UQ40 from: 1 to: 1262

```

152 AGAAGACCGCGCGCATGTCGCCAATATGCGGAGG..... 190
|||||:|||||:|||||:|||||:|||||:|||||:|||||
269 ArgArgMetGlyArgSerArgSerProAlaThrAlaArgGly 285
191 .CGGGTTGAACCGCGACGACGAG.....CGGTCAAGC 224
|||||:|||||:|||||:|||||:|||||:|||||:|||||
285 ArgSerArgThrProThrLysArgGlyHisSerArgSerArg 302
225 CGTTTTCGGAACCGCAAAATCGGTTTGAACTTGCC..... 265
|||:|||||:|||||:|||||:|||||:|||||:|||||
302 IntParArgSerArgSerArgSerAlaGlnArgTrpGlyArg 318
266 .....CCGCGTTTTCAAAAA 282
|||||:|||||:|||||:|||||:|||||:|||||:|||||
319 GlnArgArgGlyArgSerArgSerProGlnArgProGlyTrp 335
1283 CCGGAAGCATCGAACAATGTCAAGCG..... 313
|||||:|||||:|||||:|||||:|||||:|||||:|||||
335 ArgGlnThrGlnArgArgGlyArgSerArgSerAlaArgGly 352
314 .....TACAGGCTGGGACACAGCTGACGCGGCTTTGG 346

```

```

352 erHsSerArgSerProAlaThrArgGlyArgSerArgSerArg... 366
347 ACAAGGGCGAAGGGCTGCTGTCATCAGCGACATGGCGACATGAT 396
367 ThrProAlaArgArgGlyArgSerArgSerArgThrProAla... 380
397 TTGGGGGAGAGCTACATCAGCAGCAGCTTCGTTCCACCTGACGCCAT 446
380 ..... 380
447 GTTCAAGCCGGCGAATAATCAAGCGATAGCAAAATCATGCGGGGGA 496
381 .....ArgArgArgSerArgSerArgThrProThrArgArgArg... 393
497 GGGTGGCGGCAAGGCAAAACCGCGCCATACAGGAGGTCAAA 546
394 .....SerArgSerArgThrProAlaArgArgGlyArgse 405
547 CAATATCATAGGCGCTGCGCGCGGCGAGGC...AACCATCATCTGCC 593
405 rArgSer.ArgThrProAlaArgArgSerArgThrArgSerProAl 421
594 CGACACAGTCCCTTCGCCAGAGGAGCGCGGCGTGGCGGATTTT 643
422 ArgArgArgSerArgSerArgSerProAlaArgArgSerGly... 435
644 TCGGCAAACTGCATACACCATGACACTGGCGGCAAAATTGGCACAGTC 693
436 .....Args 437
694 AAAGCGCTGAACCCCTGTT.....TTTCTGCTCGGAACGCT 731
437 erArgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrPro 453
732 GCCCGACGAGCAGGCTTCGTTGCACATCCGCCCGCTCAAGGGGAAT 781
454 AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgSerGlyArg 470
782 TGAACGGCAACAAAGCCACGATGCCG 809
470 gSerArgSerArgThrProAlaArgArg 479

seq_name: sp_human:Q9UHAB

seq_documentation_block:
ID Q9UHAB PRELIMINARY: PRT: 2296 AA.
AC 01-MAY-2000 (Tremblrel. 13. Created)
DT 01-MAY-2000 (Tremblrel. 13. Last sequence update)
DI 01-DEC-2001 (Tremblrel. 19. Last annotation update)
DE SPLICING COACTIVATOR SUBUNIT SRM300.
GN SRM300.
OS Homo sapiens (Human).
OC Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI-TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20132238; PubMed=10668804;
RA Blencowe B.J., Bauren G., Eldridge A.G., Issner R., Nickerson J.A.,
RA Rosonina E., Sharp P.A.,
RT "The Srm160/300 splicing coactivator subunits.";
RL RNA 6:111-120(2000).
DR EMBL; AF201422; AAF21439.1;
DR InterPro: IPR002965; P.Rich.EXTENS.
DR PRINTS; PR01217; PRICHEXTENS.
SQ SEQUENCE 2296 AA; 251964 MW; 17C0BD4E10A9CF9 CRC64;

```

```

alignment_scores:
  Quality: 128.00      Length: 260
  Ratio: 1.153        Gaps: 11
  Percent Similarity: 42.692  Percent Identity: 26.134

```

```

alignment_block:
US-09-303-518D-571 x Q9UHAB
Align seg 1/1 to: Q9UHAB from: 1 to: 2296

152 AGAAGACCGCGCGCATCGTCGCCAATATGCGGACG..... 190
479 ArgArgMetGlyArgSerArgSerArgSerProAlaThrAlaArgGlyArgse 495
191 .CGGGTTGAACCCCGACACGACGA.....CGGTCAAGC 224
495 rArgSerArgThrProThrLysArgGlyHisSerArgSerArgSerProG 512
225 CGTTTTCGCGAAGCGCAAAATCGGTTTGAACCTTGCC..... 265
512 IntrParArgSerArgSerArgSerLacInArgTrpGlyArgSerArgSerPro 528
266 .....CCGGCTTTTCAAAA 282
529 GlnArgArgGlyArgSerArgSerArgProGlnArgProGlyTrpSerArgse 545
283 CCGGAGACATCGAACAATGTTCAAGCG..... 313
545 rArgSerArgGlnArgArgGlyArgSerArgSerArgAlaArgArgGlyArgS 562
314 .....TACACGGCTGGGAACACAGTCACAGCAGGCTTTGG 346
562 erHsSerArgSerProAlaThrArgGlyArgSerArgSerArg... 576
347 ACAAGGGGGAAGGCTGCTGTCATCAGCGCGCATCGCAGCTACGAT 396
577 ThrProAlaArgArgGlyArgSerArgSerArgThrProAla... 590
397 TTGGCGGAGCGCTACATCAGCAGCAGCTTCGTTCCACCTGACGCCAT 446
590 ..... 590
447 GTACAGCCGCGCAAAATCAAGCGATAGCAAAATATGCGAGCGGCA 496
591 .....ArgArgArgSerArgSerArgThrProThrArgArgArg... 603
497 GGTGGCGGCAAAAGCGCAAAACCGCGCACCGCATACAAAGGGGTCAAA 546
604 .....SerArgSerArgThrProAlaArgArgGlyArgse 615
547 CAATATCATAGGCGCTGCGCGCGGCGAGGC...AACCATCATCTGCC 593
615 rArgSer.ArgThrProAlaArgArgArgSerArgThrArgSerProAl 631
594 CGACACAGTCCCTTCGCCAGAGGAGCGCGGCTGTGGCGGATTTT 643
632 ArgArgArgSerArgSerArgSerProAlaArgArgSerGly... 645
644 TCGGCAAACTGCATACACCATGACACTGGCGGCAAAATTGGCACAGTC 693
646 .....Args 647
694 AAAGCGCTGAACCCCTGTT.....TTTCTGCTCGGAACGCT 731
647 erArgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrPro 663
732 GCCCGACGAGACAGCTTCGTTGCACATCCGCCCGCTCAAGGGGAAT 781
664 AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgSerGlyArg 680
782 TGAACGGCAACAAAGCCACGATGCCG 809
680 gSerArgSerArgThrProAlaArgArg 689

seq_name: sp_human:Q9UC35
seq_documentation_block:

```

ID Q9U035 PRELIMINARY; PRT: 2752 AA.
AC Q9U035;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE RNA BINDING PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohtaki S., Umeki K., Sawada Y.;
RT "Homo sapiens mRNA for RNA binding protein, complete cds.";
RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB016092; BAA83718.1; -
DR InterPro: IPR002965; P_Rich_extensn.
DR PRINTS: PRO1217; PRICHEXTENS.
SQ SEQUENCE 2752 AA; 299672 MW; 109C64F181097123 CRC64;

alignment_scores:

Quality: 128.00 Length: 260
Ratio: 1.153 Gaps: 11
Percent Similarity: 42.692 Percent Identity: 26.154

alignment_block:

US-09-303-518D-571 x Q9U035

Align seg 1/1 to: Q9U035 from: 1 to: 2752

152 AGAAGACCGCGCGCATCGCCCAATATGGCGAGC..... 190
||||| :|||: ||||| :|||:
479 ArgArgMetLylArgSerArgSerProAlaThrAlaLysArgLylArgSe 495
191 .CGGGTTTGACCCGACGACGAGA.....CGGTCAAGC 224
||| |||||:|||||
495 rArgSerArgThrProThrLysArgLylHisSerArgSerArgSerPro 512
225 CGTTTTCGCGAAGCGCAAAATGCGTTTGAACTGGCC..... 265
::: |||||:|||||
512 IntrpArgArgSerArgSerAlaGlnArgTprLylArgSerArgSerPro 528
266CGCGTTTTCAAAAA 282
529 GlnArgArgLylArgSerArgSerProGlnArgProGlyTprSerArgSe 545
283 CCGGAAGACATCGAAACATGTTCAAGCGG..... 313
:|||||:|||||: |||:
545 rArgSerThrGlnArgArgLylArgSerArgSerAlaArgArgLylArgS 562
314TACAGGCTGGGACACGTCGACGAGCGCTTGG 346
562 erHisSerArgSerProAlaThrArgLylArgSerArgSerArg..... 576
347 ACAAGGGCGAAGGCGTCTGTCATCAGCGCGACATCGCAGCTAGCAT 396
||| |||||:|||||
577 ThrProAlaArgArgLylArgSerArgSerArgThrProAla..... 590
397 TTGGGCGAGCTACATCAGCCAGCAGCTCGTCCACCTGACCGCAT 446
590 590
447 GTACAAGCGCGCGAAATCAAGCGATGACAAATTCATGAGCGCGGCA 496
|||||:|||||: ||| :|||
591ArgArgArgSerArgSerArgThrProThrArgArgArg.... 603
497 GGGTCGCGGCAAGCAAGCAAGCGCGCCACCGCATCAAGGGGTCAAA 546
604SerArgSerArgThrProAlaArgArgLylArgSe 615
547 CAATCATCAAGCGCTGCGCGCGCGAGGC...AACCATCATCTGCGC 593
:::|||| :||| ||||| ||||| :|||: |||||

615 rArgSer.ArgThrProAlaArgArgSerArgThrArgSerProAla 631
594 CGACGACGTCCCTTCCCGAGAGAGCGCGCGTGTGGCGGATTTT 643
||| ||| ||| :||| ||||| |||
632 ArgArgArgSerArgSerArgSerProAlaArgSerGly..... 645
644 TCGGCAAACTGCATACACCATGACACTGGCGCAAAATTTGGCACACGTC 693
646..... |||:
694 AAAGCGGTGAACCCCTGTT.....TTTCTGCTGGACGCGCT 731
:|||| :|||: |||||
647 erArgSerArgThrProAlaArgArgLylArgSerArgSerArgThrPro 663
732 GCCCGACGACAAAGCGTTCGTGTCACATCGCCCGTCCAAAGGGAAT 781
||||||| ||| ||| ||| ||| |||
664 AlaArgArgLylArgSerArgSerArgThrProAlaArgSerGlyAr 680
782 TGAAGGCAACAAAGCCACGATGCCGC 809
:||||:||||| |||
680 gSerArgSerArgThrProAlaArgArg 689

